

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES  
DER FAKULTÄT FÜR BIOLOGIE  
DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

# TALE NUCLEASES AS ANTIVIRAL FACTORS



**MARKUS JULIUS PRÄGERT**  
MÜNCHEN 2013

Dissertation eingereicht am 01. August 2013

Erstgutachter: Prof. Dr. Dirk Eick

Zweitgutachter: Prof. Dr. Thomas Lahaye

Tag der mündlichen Prüfung: 08. Januar 2014

# ERKLÄRUNG

Hiermit erkläre ich, dass die vorliegende Arbeit mit dem Titel

„TALE NUCLEASES AS ANTIVIRAL FACTORS“

von mir selbstständig und ohne unerlaubte Hilfsmittel angefertigt wurde, und ich mich dabei nur der ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe. Die Arbeit wurde weder in der jetzigen noch in einer abgewandelten Form einer anderen Prüfungskommission vorgelegt.

München, 01. August 2013

Markus Julius Prägert





# Content

<b>1. Introduction</b>	<b>1</b>
<b>1.1 Prologue</b>	<b>1</b>
<b>1.2 ZFN</b>	<b>2</b>
1.2.1 Zinc finger proteins	2
1.2.2 Zinc finger nucleases	4
<b>1.3 TALEN</b>	<b>6</b>
1.3.1 Transcription activator-like effectors (TALEs)	6
1.3.2 Transcription activator-like effector nucleases (TALENs)	9
<b>1.4 Genome engineering using custom-tailored nucleases</b>	<b>10</b>
1.4.1 Non-homologous DNA end-joining (NHEJ)	11
1.4.2 Homologous recombination (HR)	12
1.4.3 Genome engineering	13
<b>1.5 Murine gammaherpesvirus 68 (MHV-68)</b>	<b>16</b>
<b>1.6 Aims</b>	<b>17</b>
<b>2. Materials</b>	<b>19</b>
<b>2.1 Oligonucleotides</b>	<b>19</b>
<b>2.2 Plasmids</b>	<b>19</b>
<b>2.3 Antibodies</b>	<b>21</b>
<b>2.4 Bacterial strains</b>	<b>21</b>
<b>2.5 Eukaryotic cell lines</b>	<b>22</b>
<b>2.6 Cell culture media and additives</b>	<b>22</b>
2.6.1 Media for the cultivation of prokaryotic cells	22
2.6.2 Media for the cultivation of eukaryotic cells	23
<b>2.7 Chemicals and Enzymes</b>	<b>23</b>
<b>2.8 Buffers and solutions</b>	<b>24</b>

Content	II
<b>2.9 Commercial kits</b>	<b>26</b>
<b>2.10 Devices and consumables</b>	<b>27</b>
<b>2.11 Software</b>	<b>28</b>
<b>3. Methods</b>	<b>29</b>
<b>3.1 Bacterial culture</b>	<b>29</b>
3.1.1 Growth and storage of bacterial cultures	29
3.1.2 Transformation of bacteria	29
<b>3.2 Eukaryotic cell culture</b>	<b>30</b>
3.2.1 Cell culture conditions	30
3.2.2 Storage of eukaryotic cells	31
3.2.3 Transfection of HEK293 cells	32
3.2.4 Electroporation of U-2 OS / K562 cells	32
3.2.5 Generation of stable cell lines	33
3.2.6 Staining of cells	34
3.2.7 Flow cytometry	35
3.2.8 Generation of clonal cell lines	35
3.2.9 Cell separation using MACS beads	35
3.2.10 Cell sorting using BD FACSAria III	36
3.2.11 Measurement of cell proliferation using the Xcelligence system	36
<b>3.3 Virological techniques</b>	<b>37</b>
3.3.1 Infection of HEK293 derived cell lines with MHV-68	37
3.3.2 Generating viral supernatants	37
3.3.3 Determination of viral titer	37
<b>3.4 Nucleic acid techniques</b>	<b>38</b>
3.4.1 DNA purification from E.coli	38
3.4.2 DNA purification from eukaryotic cells	39
3.4.3 DNA isolation from agarose gels	39
3.4.4 DNA isolation from SDS gels	39
3.4.5 Dephosphorylation and ligation	39
3.4.6 Polymerase chain reaction (PCR)	39
3.4.7 Quantitative real time PCR (qPCR)	40
3.4.8 Mutagenesis PCR	41
3.4.9 Cut-ligation of TALEN	41
<b>3.5 Protein analysis techniques</b>	<b>42</b>
3.5.1 <i>In vitro</i> transcription and translation	42
3.5.2 Preparation of cell extracts	43

Content	III
3.5.3 Strep-Tactin affinity chromatography	43
3.5.4 Electrophoretic mobility shift assay (EMSA)	44
3.5.5 Determination of the Equilibrium Dissociation Constant ( $K_d$ -value)	45
3.5.6 SDS-PAGE	46
3.5.7 Western blot	46
<b>3.6 <i>In vitro</i> cut assay</b>	<b>47</b>
<b>3.7 <i>In vivo</i> cut assay</b>	<b>47</b>
<b>3.8 Measurement of the DNA damage response (DDR)</b>	<b>48</b>
<b><u>4. Results I – Biochemical characterization of DNA binding attributes of ZFNs and TALEs</u></b>	<b><u>49</u></b>
<b>4.1 Generation of <i>gfp</i>-specific TALENs</b>	<b>49</b>
<b>4.2 Overview of TALE/Ns and ZFNs used in this thesis</b>	<b>52</b>
<b>4.3 Protein purification with HEK293 cell lysates</b>	<b>54</b>
<b>4.4 Electrophoretic mobility shift assays with ZFNs</b>	<b>55</b>
<b>4.5 Electrophoretic mobility shift assays with TALEs</b>	<b>58</b>
<b>4.6 Electrophoretic mobility shift assays with the c-Jun/c-Fos dimer</b>	<b>60</b>
<b><u>5. Results II – TALENs as antiviral factors</u></b>	<b><u>62</u></b>
<b>5.1 Analysis of the cleavage activity</b>	<b>62</b>
5.1.1 <i>In vitro</i> cleavage assay	62
5.1.2 <i>In vivo</i> cleavage assay	64
<b>5.2 Establishment and characterization of an inducible expression system</b>	<b>66</b>
5.2.1 First induction and expression analysis	66
5.2.2 Comparison of different doxycycline concentrations	68
5.2.3 Single cell clones and their dox inductivity	69
5.2.4 Correlation between CD2 expression and induction rate	71
5.2.5 Kinetics of dox-induced gene expression	73
<b>5.3 Infection of TALEN expressing cells with MHV-68-gfp</b>	<b>74</b>
5.3.1 Knockout of viral GFP	74
5.3.2 Analysis of TALEN-mediated cleavage of the viral <i>gfp</i> sequence by qPCR	77
<b>5.4 Interference of TALEN expression with viral <i>de novo</i> synthesis</b>	<b>79</b>
5.4.1 MHV-68-gfp <i>de novo</i> synthesis in TALEN or mCherry expressing cells	79
5.4.2 MHV-68-gfp <i>de novo</i> synthesis in highly induced TALEN or mCherry expressing cells	81
5.4.3 A comparison of MHV-68 and MHV-68-gfp <i>de novo</i> synthesis in TALEN expressing HEK293 cells	84
<b>5.5 Generation of TALENs with mutated FokI domains</b>	<b>87</b>
5.5.1 Overview of TALENs with mutated FokI domains	87

Content	IV
5.5.2 <i>In vitro</i> cleavage assay comparing TALENs with mutated FokI domains	88
5.5.3 <i>In vivo</i> cleavage assay comparing TALENs with mutated FokI domains	90
<b>5.6 Analysis of TALEN-mediated genotoxicity</b>	<b>92</b>
5.6.1 DNA damage response in cells expressing different TALEN pairs	92
5.6.2 The kinetic of the DNA damage response as a function of TALEN expression	94
<b>5.7 Application of TALENs with mutated FokI domains</b>	<b>96</b>
5.7.1 Establishment and analysis of a cell line expressing TALENs with mutated FokI domains	96
5.7.2 MHV-68 and MHV-68-gfp <i>de novo</i> synthesis in cells expressing TALENs with mutated FokI domains	98
<b>5.8 Effect of induced TALEN expression on cellular proliferation</b>	<b>100</b>
5.8.1 Establishment of two additional inducible cell lines	100
5.8.2 Proliferation analysis of inducible cell lines	102
<b>6. Discussion</b>	<b>106</b>
<b>6.1 Scope of this project</b>	<b>106</b>
<b>6.2 Biochemical analysis of ZFN and TALE DNA binding</b>	<b>108</b>
6.2.1 ZFN	108
6.2.2 TALE	110
<b>6.3 TALENs as antiviral factors</b>	<b>111</b>
6.3.1 Generation of <i>gfp</i> -specific TALENs	111
6.3.2 Cleavage of viral genomes	112
6.3.3 Inhibition of viral replication and genotoxic effects	114
<b>6.4 Possible applications</b>	<b>118</b>
<b>6.5 Open questions and outlook</b>	<b>120</b>
<b>7. Summary</b>	<b>123</b>
<b>8. Abbreviations</b>	<b>125</b>
<b>9. Literature</b>	<b>128</b>
<b>10. Appendix</b>	<b>138</b>
<b>10.1 Oligonucleotides</b>	<b>138</b>
10.1.1 Mutagenesis PCR primers	138
10.1.2 PCR primers	139
10.1.3 qPCR primers	139
10.1.4 EMSA DNA templates	139
<b>10.2 Amino acid sequences</b>	<b>141</b>

Content	V
10.2.1 Zinc finger nucleases	141
10.2.2 TAL effector / nucleases	141
<b>11. Publications</b>	<b>145</b>

---

# 1. Introduction

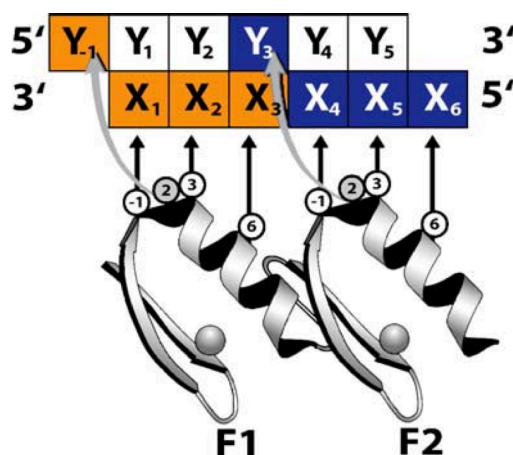
## 1.1 Prologue

Deoxyribonucleic acid (DNA) is one of the three essential macromolecules in living organisms. The genetic information of an organism is encoded by the nucleotides guanine (G), adenine (A), thymine (T), and cytosine (C). Functional units encoded by the sequence of nucleotide residues are called genes and determine the composition, i.e. primary sequence information of the other two essential macromolecules, ribonucleic acids (RNA) and proteins. Certain proteins interact directly with DNA. These proteins fulfill two essential functions, which are carried out by protein domains. In essence one domain achieves binding of the protein to DNA. In many cases this domain can “read” the sequence of a stretch of nucleotide residues and therefore mediates the sequence-specific binding of the protein to a certain sequence motif of a defined length. The second domain is called the effector domain and determines the task or function(s) that are executed by the protein. These can be the recruitment of other proteins to DNA, regulation of gene expression or the cleavage of DNA. Recently, researchers gained insight into how sequence-specific DNA binding domains of a certain class of transcription factors can be modified to alter their target sequence redirecting them to predetermined sequence motifs. This knowledge provided the basis for constructing DNA binding domains with novel specificities to almost any sequence composition. This recent strategy led to the development of new and important tools in molecular biology, opening the door for further biotechnological applications. One new emerging field is the construction of proteins, which cleave DNA sequence specifically at a predetermined site. These so called “designer nucleases” were constructed by fusing nuclease domains, which cleave DNA, to customized DNA binding domains.

## 1.2 ZFN

### 1.2.1 Zinc finger proteins

In 1985, the zinc finger motif was described for the first time (BROWN *et al.* 1985; MILLER *et al.* 1985). Both groups described the structure of the transcription factor TFIIIA in *Xenopus laevis* and postulated that its DNA binding domain consists of nine tandem



**Fig. 1.1 Model of two Cys<sub>2</sub>His<sub>2</sub> zinc fingers binding to DNA.** The  $\beta\beta\alpha$  structure is stabilized by a zinc ion (grey sphere). Zinc finger F1 and F2 bind to the major groove of the DNA with amino acids -1, 3 and 6. Amino acid 2 targets one base pair (bp) on the complementary strand of the DNA extending the recognized sequence. Figure adopted from Papworth *et al.* 2006.

repeats with a length of 30 amino acids, each containing two cysteine and two histidine residues. The authors assumed that these residues are coordinated by a central zinc ion. The Cys<sub>2</sub>His<sub>2</sub> zinc finger proteins were the founding members of the zinc finger superfamily. It became evident that zinc fingers are one of the most abundant and important DNA binding domains in eukaryotes (PELLEGRINO and BERG 1991; JACOBS 1992), which are able to bind to DNA, RNA (SHAstry 1996) and interact with other proteins (MACKAY and CROSSLEY 1998). Crystal structures confirmed the modular structure with each finger containing a short two-stranded antiparallel  $\beta$  sheet and an  $\alpha$  helix

(Fig. 1.1). The zinc ion is coordinated by two conserved cysteine residues contributed by the  $\beta$  sheet and two histidine residues contributed by the  $\alpha$  helix (Fig. 1.1; PAVLETICH and PABO 1991; ELROD-ERICKSON *et al.* 1996). The stability of the  $\beta\beta\alpha$  architecture is derived from the crosslinking provided by the zinc ion between the  $\beta$  sheet and the  $\alpha$  helix (FRANKEL *et al.* 1987). In transcription factors zinc fingers are typically arranged in an array of three or four units each consisting of a zinc atom coordinating a single zinc finger (JAMIESON *et al.* 2003). These composite zinc finger transcription factors are by far the most abundant class of DNA binding domains in human transcription factors (TUPLER *et al.* 2001). The structure revealed that one zinc finger targets a contiguous three base pairs (bp) DNA sequence with the side chains of the N-terminal amino acids at the

positions -1, 3 and 6 (numbers are with respect to the start of the helix) of the  $\alpha$  helix (Fig. 1.1, PAVLETICH and PABO 1991). Besides the three base pairs, which a single zinc finger contacts it was shown that the amino acid at position 2 targets one bp on the complementary DNA strand. This extends the recognized nucleotide sequence and restricts the independent DNA binding of zinc finger modules (ISALAN *et al.* 1997).

Several groups showed that changes of key residues in the  $\alpha$  helix can lead to zinc fingers with novel DNA target sequences (JAMIESON *et al.* 1994; REBAR and PABO 1994; WOLFE *et al.* 2000). These findings were exploited to generate designed zinc finger modules with new specificities (REBAR *et al.* 1996; JOUNG *et al.* 2000; WOLFE *et al.* 2000). Since zinc fingers do not bind to DNA context-independently but interfere with adjacent zinc fingers the generation of new DNA binding domains via modular assembly of different zinc finger modules is appealing in its simplicity. But modular assembly showed only low efficacy (RAMIREZ *et al.* 2008) yielding also zinc finger nucleases with low activities and high toxicities (CORNU *et al.* 2008; PRUETT-MILLER *et al.* 2008). Therefore, synthetic zinc fingers require context sensitive strategies for their generation, which identify zinc finger module combinations that are compatible and do work well together. These methods are labor-intensive because they make use of selection methods like phage display or the bacterial cell-based two-hybrid system. They deliver multi-finger domains with high DNA binding affinities and specificities but need large randomized libraries and require expertise (CATHOMEN and JOUNG 2008). A more rapid method to generate highly active zinc fingers efficiently is the publicly available Oligomerized Pool Engineering (OPEN), which is based on preselected zinc finger pools (MAEDER *et al.* 2008).

This knowledge offered an attractive framework for the generation of new DNA modifying enzymes on the basis of DNA binding domains derived from engineered zinc finger proteins with novel specificities. Engineered zinc finger domains with desired DNA binding specificities were equipped with various effector domains like the VP16 activation domain from herpes simplex virus or the KRAB repression domain (BEERLI *et al.* 1998) to generate designer transcription factors altering the transcription of target genes (BLANCAFORT *et al.* 2004). Zinc fingers were also combined with DNA methyltransferases (Xu and Bestor 1997) leading to target-specific methylation of cytosines (SMITH and FORD 2007), integrases leading to a directed integration of



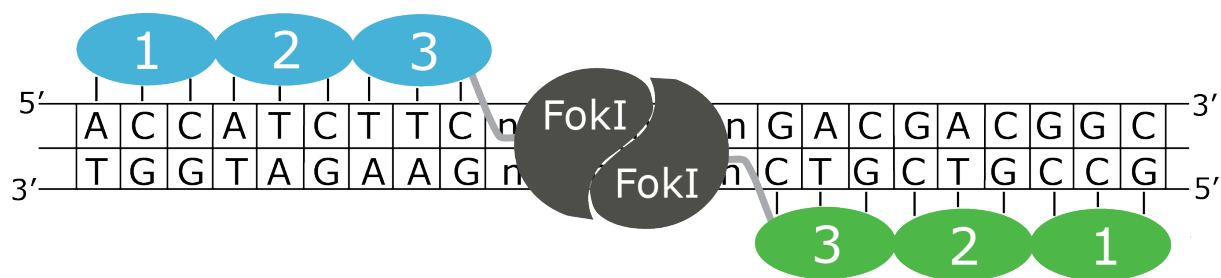
retroviral DNA to predetermined sites (BUSHMAN and MILLER 1997) and also the cleavage domain of the FokI endonuclease to create novel, site-specific endonucleases (KIM *et al.* 1996).

### 1.2.2 Zinc finger nucleases

In 1996, the Chandrasegaran group opened a new field of research when they fused the cleavage domain of the FokI endonuclease with DNA binding zinc finger proteins (KIM *et al.* 1996). In an earlier publication they could show that the FokI restriction enzyme consists of two domains, a recognition domain, which mediates sequence-specific binding to a DNA motif and a nonspecific cleavage domain (LI *et al.* 1992) fused by a linker domain (LI and CHANDRASEGARAN 1993). This modular structure provided the possibility to combine the cleavage domain of FokI with a zinc finger DNA binding domain, generating restriction enzymes with novel sequence specificities, the zinc finger nucleases (ZFNs). The cleavage domain of FokI has to dimerize to become catalytically active (BITINAITE *et al.* 1998) determining the general conditions for the construction of ZFNs. Two inverted zinc finger binding sites one at the top and bottom strand of DNA separated by a spacer of six bp are necessary and sufficient to bring the FokI domains in close proximity and confer dimerization, which results in efficient DNA cleavage (SMITH *et al.* 2000). One ZFN usually harbors three to four zinc fingers, which make contact with nine to twelve nucleotides of the target sequence. A ZFN pair therefore recognizes a specific sequence of 18 to 24 bp separated by a six bp spacer (MILLER *et al.* 2007; SZCZEPEK *et al.* 2007). The FokI cleavage domain is connected to the zinc finger DNA binding domain by a short inter-domain linker, which can vary in size and length (HANDEL *et al.* 2009). The ZFN pair employed in this thesis is depicted in Figure 1.2. ZFNs were successfully used in several different organisms and cell lines (PORTEUS *et al.* 2003; ALWIN *et al.* 2005; URNOV *et al.* 2005; WRIGHT *et al.* 2005; BEUMER *et al.* 2006; MAEDER *et al.* 2008; MENG *et al.* 2008; KIM *et al.* 2009).

It became apparent that the specificity of DNA binding is a crucial parameter in the application of ZFNs. A lack of specificity was shown to result in ZFN-mediated genotoxicity, which derives from induced DNA double-strand breaks at off-target sites

(PORTEUS and BALTIMORE 2003; ALWIN *et al.* 2005). Considerable effort has been put into research to overcome this burden. It was shown that the cytotoxicity induced by ZFNs varies with the length of the inter-domain linker. The optimal results regarding activity and low cytotoxicity were obtained with the combination of a six bp spacer and a four amino acid inter-domain linker (HANDEL *et al.* 2009).



**Fig. 1.2** Zinc finger nucleases (ZFN) consist of a DNA binding domain (blue/green), the cleavage domain of the FokI restriction enzyme (dark grey) and a linker fusing the two domains together (light grey). The scheme shows the ZFN pair used in this thesis. The DNA binding domains consist of three zinc fingers each (blue/green). Each zinc finger recognizes three bp (black vertical lines) resulting in a nine bp target sequence for each ZFN separated by a six bp spacer. Both ZFN together recognize a target sequence of 18 bp as shown.

The second focus of consideration to reduce toxicity while maintaining a high activity was the dimerization of the FokI domain. The FokI restriction enzyme forms homodimers of two identical monomers that recognize and bind to a palindromic DNA sequence. Here, two different molecules, ZFN 1 and ZFN 2, must form dimers in order to direct it to the correct DNA sequence motif. Since the FokI nuclease domains are identical in ZFN 1 and ZFN 2 dimers can form that consist of two ZFN 1 or two ZFN 2 molecules (Fig. 1.2 blue or green). The generation of obligate heterodimerizing FokI domains solved this problem and increased specificity of DNA cleavage (MILLER *et al.* 2007; SZCZEPEK *et al.* 2007; DOYON *et al.* 2011). The third point that was addressed was the DNA binding specificity of the zinc finger domains. It was shown that a high binding specificity results in a high nuclease activity at the target site and a low ZFN-mediated toxicity (CORNU *et al.* 2008).

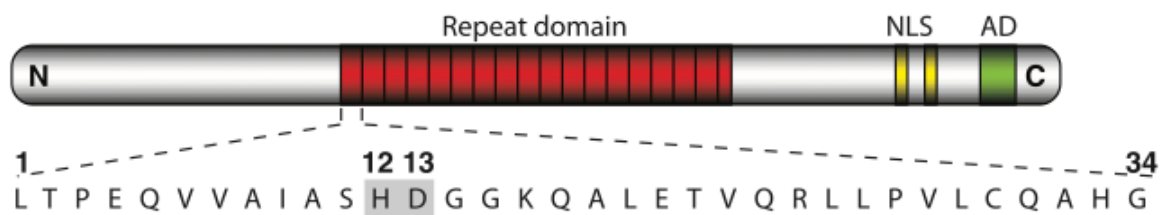
## 1.3 TALEN

### 1.3.1 Transcription activator-like effectors (TALEs)

Plants are exposed to a high number of potential microbial pathogens but overcome this threat with their two-branched innate immune system (JONES and DANGL 2006). The first branch is based on the recognition of pathogen/microbial associated patterns (PAMPs or MAMPs) through pattern recognition receptors (PRR), which recognize specific conserved epitopes like a stretch of 22 amino acids of the microbial flagellin. The so-called PAMP-triggered immunity (PTI) releases defense responses in a cell-autonomous manner. If a pathogen cannot overcome the PTI its proliferation is successfully prevented (BOLLER and HE 2009). Bacteria can break this barrier by secreting effectors via their type III secretion system into plant cells. These effector molecules are designed to manipulate the plant immunity and inhibit PTI resulting in effector triggered susceptibility (GOHRE and ROBATZEK 2008). Plants in turn use the second branch of their immune system to perceive such effectors. The plant encodes certain proteins that recognize pathogen effectors mediating resistance to the pathogen by inducing effector-triggered immunity (BELKHADIR *et al.* 2004; BOGDANOVE *et al.* 2010).

The phytopathogenic bacteria of the genus *Xanthomonas* infect a wide range of host plants and are responsible for important crop plant diseases like bacterial blight of rice caused by *Xanthomonas oryzae*. The bacteria inject the transcription activator-like effector (TAL effector or TALE) into plant cells where it acts as a transcription factor (KAY and BONAS 2009) and activates host susceptibility genes. The activation of host susceptibility genes facilitates bacterial colonization and spread resulting in disease promotion (BOGDANOVE *et al.* 2010). The N-terminal part of the *Xanthomonas* TAL effector AvrBs3 was shown to be important for its translocation into the host cell via the bacterial type III secretion system (SZUREK *et al.* 2002). The C-terminal part harbors an acidic transcription activation domain and two nuclear localization signals, which were shown to interact with importin  $\alpha$  of the nuclear import machinery shuttling the protein to the nucleus (SZUREK *et al.* 2001).

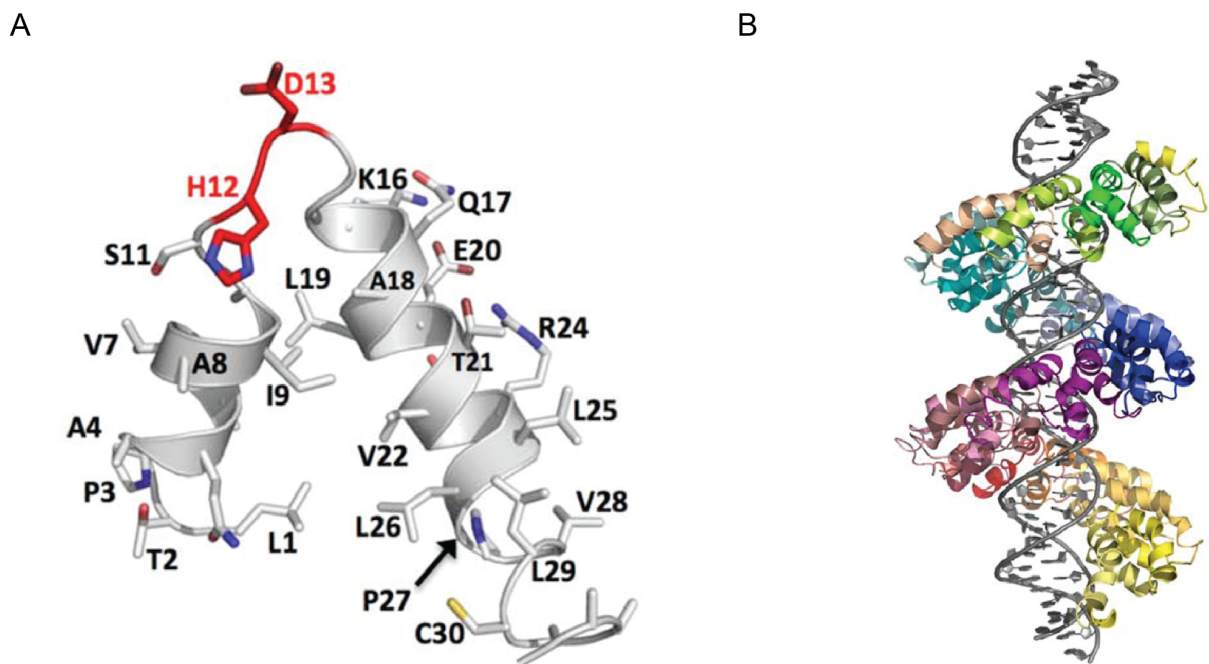
The central domain of AvrBs3 consists of 17.5 repeats (BONAS *et al.* 1989) each consisting of 34 amino acids. An exception is the terminal half repeat, which consists of 20 amino acids, only. The amino acid sequences of these repeats are highly conserved except amino acid 12 and 13, which are variable (Fig. 1.3; SCHORNACK *et al.* 2006). The central domain confers sequence specific binding of AvrBs3 to DNA (KAY *et al.* 2007) but the nature of DNA binding remained enigmatic until 2009 when two groups deciphered the underlying code (BOCH *et al.* 2009; MOSCOU and BOGDANOVE 2009).



**Fig. 1.3** The scheme shows the TAL effector AvrBs3. The N-terminal domain is necessary for translocation of AvrBs3 to the host cell via the type III secretion system. The C-terminal domain harbors an acidic transcription activation domain (AD) and two nuclear localization signals (NLS). The central repeat domain confers sequence specific binding to host cell DNA. Each of the 17 repeats consists of 34 amino acids, the last half repeat consists of 20 amino acids. The amino acid sequence of the first repeat is depicted below the scheme. The 34 amino acids are highly conserved, except amino acid 12 and 13, the so-called repeat-variable di-residue (RVD), specifying the DNA binding in a one-to-one correspondence between one RVD and a single nucleotide residue. This Figure was adopted from Boch *et al.* 2009.

The two groups could show that one repeat unit of the central repeat domain binds one specific DNA base pair in a one-to-one correspondence. As a consequence, the DNA binding domain of AvrBs3 with its 17.5 repeats targets an 18 bp contiguous nucleotide sequence. Each single repeat acts independently and makes contact to a single nucleotide. The specificity of a single repeat is defined by the variable amino acids 12 and 13, the repeat-variable di-residue (RVD). The two groups showed that RVDs consisting of the amino acids HD, NI and NG have a strong preference for the DNA nucleotides C, A and T, respectively. G is targeted with the RVD NN showing also a substantial binding to A. An increase in specificity was suggested with the RVD NK, showing a much stronger preference for G residues (MORBITZER *et al.* 2010; MILLER *et al.* 2011) but also reducing TALE-mediated gene-transactivation (HUANG *et al.* 2011) due to a lower binding affinity for its target DNA sequence (CHRISTIAN *et al.* 2012). A repeat containing the amino acids NS at position 12 and 13 can target all four nucleotides. This simple two-letter code allows to predict the target sequence of a TAL effector or the engineering of a customized repeat array with a novel DNA sequence specificity. Published crystal structures revealed that each repeat comprises two left-handed

helices connected by a short loop containing the RVD (Fig. 1.4 A). The first RVD residue stabilizes the protein backbone; the second RVD residue makes the base-specific contact to DNA. The repeat domain forms a right-handed superhelical structure wrapped around the DNA major groove (Fig. 1.4 B; DENG *et al.* 2012; MAK *et al.* 2012). In addition, the N-terminus of the TAL effector features four additional degenerated repeats, which are essential for DNA binding. Whether these repeats contribute to DNA binding specificity or not is controversially discussed (GAO *et al.* 2012; MAK *et al.* 2012).



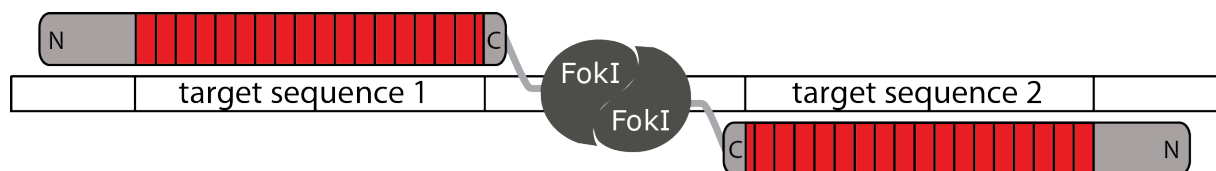
**Fig. 1.4** Crystal structure of the TAL effector PthXo1. **A.** A single repeat consists of two left-handed helices ranging from amino acid 3-11 and 14-33. The RVD is located between the two helices, consisting of amino acid 12 and 13. Amino acid 12 stabilizes the protein backbone; amino acid 13 makes the base-specific contact to DNA. **B.** The repeat domain forms a right-handed superhelical structure wrapped around the DNA major groove. The PthXo1 repeat domain consists of 23.5 repeats with additional two degenerated repeats (0 and -1) preceding the repeat domain (dark green and green) and specifying the T, which precedes the repeat domain specified sequence. Figure adopted from Mak *et al.* 2012.

Several groups have shown that the repeats of the TAL effector DNA binding domain work independent of its context (MORBITZER *et al.* 2010; MILLER *et al.* 2011; ZHANG *et al.* 2011) very much in contrast to ZFN where modules can interfere with neighboring modules (ISALAN *et al.* 1997). This advantageous feature of TAL effector repeats permits the modular assembly of repeats enabling the efficient and rapid generation of DNA binding domains with novel specificities (CERMAK *et al.* 2011; MORBITZER *et al.* 2011; REYON *et al.* 2012). In fact, recent synthetic TAL effectors generated with new DNA

binding specificities were shown to modulate transcription efficiently (MORBITZER *et al.* 2010; ZHANG *et al.* 2011).

### 1.3.2 Transcription activator-like effector nucleases (TALENs)

ZFNs consist of a zinc finger DNA binding domain and the cleavage domain of the FokI restriction enzyme (KIM *et al.* 1996). Similar to this structure TALE nucleases (TALENs) were constructed by replacing the transcriptional activation domain of the TAL effector with the endonuclease cleavage domain of the FokI nuclease, which lacks nucleotide specificity (CHRISTIAN *et al.* 2010; LI *et al.* 2011; MAHFOUZ *et al.* 2011). The FokI cleavage domain has to dimerize to become catalytically active (BITINAITE *et al.* 1998). To allow FokI dimerization two TALEN have to bind to DNA in a tale-to-tale conformation with a proper spacer length to bring the FokI domains in close proximity and thereby confer dimerization (Fig 1.5). The constructed TALENs showed comparable cleavage activity to ZFNs but with a longer optimal spacer length of about 15 bp in contrast to ZFNs, which operate with a spacer of about 6 nucleotides.



**Fig. 1.5** TAL effector nucleases (TALENs) were constructed by replacing the transcriptional activation domain of TAL effectors with the endonuclease domain of the FokI restriction enzyme. Since FokI is active as a dimer, only, two TALENs have to be designed in a way that they bind their target sequences in a tale-to-tale conformation separated by a spacer of appropriate length. The sequence specific binding of the TALENs to their target sequences brings the FokI domains into close proximity, confers dimerization and DNA cleavage.

The TAL effector serves as a DNA binding domain, only. Therefore different truncated versions of the protein were constructed and compared regarding their ability to bind to and cleave DNA when fused to FokI cleavage domains (MUSSOLINO *et al.* 2011). Upstream of the repeat domain 100 amino acids of the N-terminus were necessary for efficient DNA binding and cleavage. In contrast, the first 135 amino acids of the N-terminus are exclusively dedicated to protein transport into plant cells. This part can be removed without affecting DNA binding (SZUREK *et al.* 2002; MILLER *et al.* 2011). The residues at

the C-terminus of the repeat domain were also found dispensable for efficient DNA binding.

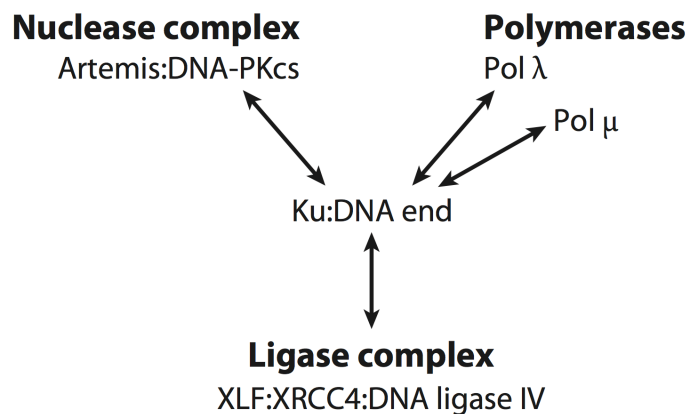
In this thesis I employed truncated TALENs, which lack the first 135 amino acids of the N-terminus and the last 261 amino acids of the C-terminus. TALENs based on this truncated scaffold showed optimal cleavage activity with a spacer length of 12 bp in contrast to full length TALEN showing an optimal cleavage activity with a spacer length of 15 bp. This difference indicates that the optimal spacer length correlates with the length of the C-terminus. TALENs based on the truncated scaffold showed cleavage activity comparable to a ZFN control with significantly reduced cytotoxicity (MUSSOLINO *et al.* 2011).

## **1.4 Genome engineering using custom-tailored nucleases**

DNA double-strand breaks (DSB) are one of the most deleterious forms of DNA lesions. They can be caused by ionizing radiation, X-rays, free radicals, chemicals, DNA replication across a nick, or failures of nuclear enzymes like type II topoisomerase, which has endonucleolytic activity (LIEBER 2010). Failure to repair DSBs can lead to chromosome loss, chromosomal translocation, apoptosis or carcinogenesis (MORRISON and TAKEDA 2000; HOEIJMAKERS 2001; LIEBER *et al.* 2003). Since DNA DSBs do have a destructive potential, endogeneous repair pathways have evolved that are highly conserved between diverse cells and organisms (HOEIJMAKERS 2001). Repair of DSBs can occur by one of the two DNA repair pathways: Non-homologous DNA end-joining (NHEJ) or homologous recombination (HR). Custom-made nucleases with novel target sequences introduce sequence-specific DSBs at will. Their repair relies on the intrinsic DNA repair pathways often in conjunction with exogenously added DNA repair templates that foster homologous recombination events at the site of DSBs, introducing novel DNA sequences. This system makes TALENs valuable tools for precise genome engineering.

### 1.4.1 Non-homologous DNA end-joining (NHEJ)

There are an estimated ten DSBs per day per cell in early passage human fibroblasts (LIEBER 2010). Non-homologous DNA end-joining (NHEJ) is the main repair mechanism of DNA DSBs during G0, G1 and early S phases when sister chromatids are not available for homologous recombination. NHEJ is an error prone repair mechanism because on both ends of the DNA break a few nucleotides are lost making NHEJ an imperfect process from the standpoint of preserving genetic information (LIEBER 1999). Typically, at least one and often up to 10-20 nucleotides are deleted at each DNA end (LIEBER and KARANJAWALA 2004).



**Fig. 1.6 DNA repair by NHEJ.** When a DSB occurs the DNA ends are first bound by Ku. The repair of the lesion is an orchestration of (i) nuclease activity, resecting damaged DNA; (ii) polymerase activity, filling in new DNA and (iii) ligase activity restoring the formal integrity of the DNA strands. Figure adopted from Lieber 2010.

When a DSB occurs Ku protein (a heterodimer of Ku70 and Ku86) binds to both DNA ends. The Ku:DNA complex serves as a platform at which the nuclease, polymerase and ligase of NHEJ can dock (LIEBER 2008). The nuclease complex Artemis:DNA-PKcs (DNA-dependent protein kinase) endonucleolytically cuts the damaged DNA overhangs (MA *et al.* 2005). In this complex, DNA-PKcs phosphorylates Artemis whereby Artemis acquires endonucleolytic activity (MA *et al.* 2002). Additionally, end processing by pol μ and pol λ occurs filling in the gaps and extending the single-stranded 3' and at 5' overhangs (LIEBER *et al.* 2004; MA *et al.* 2004). The ligase complex XLF:XRCC4:DNA ligase IV completes the joining reaction (GRAWUNDER *et al.* 1997). In this complex XRCC4 stimulates the activity of DNA ligase IV (GRAWUNDER *et al.* 1997) and XLF stimulates the joining of incompatible ends (GU *et al.* 2007). These processes can occur flexibly, repeatedly and independently on both DNA ends. This mechanistic flexibility is essential



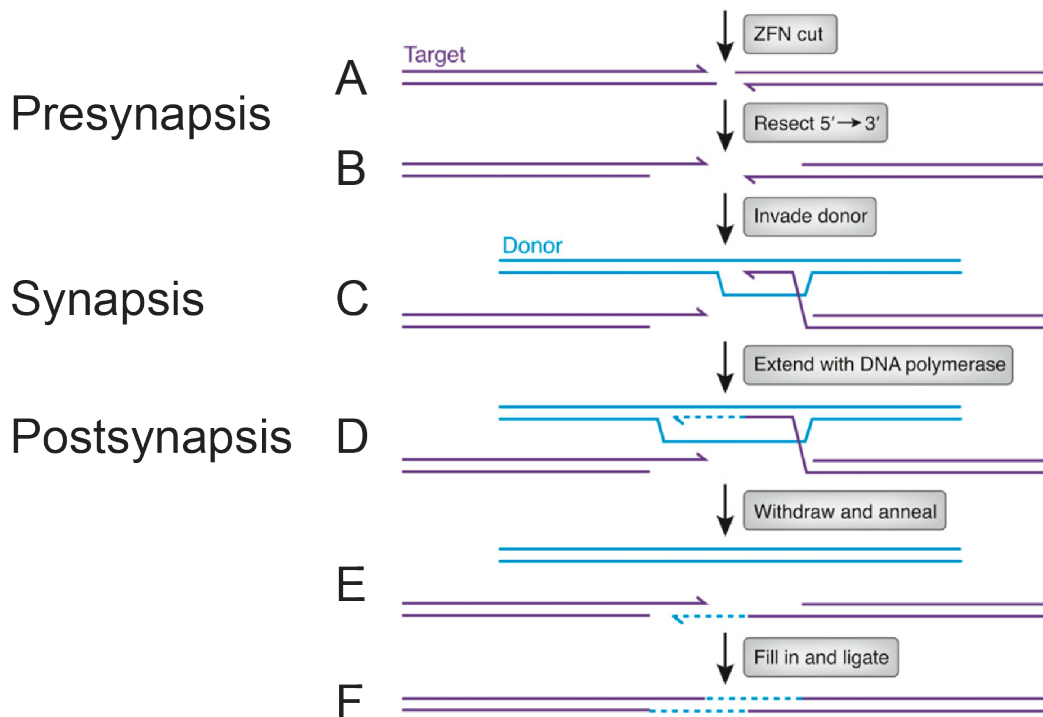
to permit NHEJ for every possible DSB and causes different outcomes after DNA DSB repair (LIEBER 2010). Fig. 1.6 summarizes the interaction between the different enzymes.

### 1.4.2 Homologous recombination (HR)

Homologous recombination is the dominating repair pathway of DNA DSBs in late S and G2 phase of the cell cycle, when the cell can provide the optimally positioned sister chromatid as a homologous template for repair, which is required for HR in contrast to NHEJ (TAKATA *et al.* 1998; LIEBER *et al.* 2003). Since HR uses a DNA template to copy the missing information typically no information is lost. HR recombination provides high-fidelity repair of complex DNA damages including DNA gaps, DNA DSBs and DNA interstrand crosslinks. Three different sub-pathways of HR are proposed (i) double Holliday junction pathway (dHJ) (SZOSTAK *et al.* 1983), (ii) synthesis-dependent strand annealing (SDSA) (NASSIF *et al.* 1994) and (iii) break-induced replication (MALKOVA *et al.* 1996). The SDSA pathway inherently avoids crossovers, which reduces the potential of genomic rearrangements in contrast to the dHJ pathway, making it the preferred pathway during mitosis. During meiosis the dHJ pathway is used to form crossovers. The break-induced replication is adopted if the second end of a DSB is lost (KREJCI *et al.* 2012). Here the focus is on synthesis-dependent strand annealing, which appears to be the predominant pathway for DSB repair in somatic cells (PAQUES and HABER 1999; HEYER *et al.* 2010).

The homologous recombination mechanism consists of three stages: presynapsis, synapsis, and postsynapsis (Fig. 1.7; LI and HEYER 2008). During presynapsis DSBs are recognized and processed to a 3'-OH ending single-strand (Fig 1.7 A, B). The 5'-3' resection takes place in two steps. In vertebrate cells the first action of the MRN complex, consisting of MRE11 nuclease, RAD50 DNA binding protein and NBS1, initiates the resection by an endonuclease mechanism that removes the 5' strand in increments of around 50-100 nucleotides. These short ssDNA overhangs are then substrates for further 5'-3' resection by the helicase BLM and the nucleases EXO1 and DNA2 (MIMITOU and SYMINGTON 2008). In the second step (synapsis) the generated ssDNA overhang is bound by replication protein A (RPA), which eliminates secondary structures. This

complex is the basis for subsequent binding of RAD51, which replaces RPA (HEYER *et al.* 2010). RAD51 bound to ssDNA, the RAD51 filament, catalyzes homology search and strand invasion generating the D-loop (Fig 1.7 C; LI and HEYER 2008). The invaded strand is extended by a DNA polymerase (Fig. 1.7 D) and re-anneals to the second 3'-single strand of the DSB through complementary base pairing, catalyzed by RAD52 (Fig. 1.7 E). DNA synthesis fills in the remaining single-stranded gaps (Fig. 1.7 F).

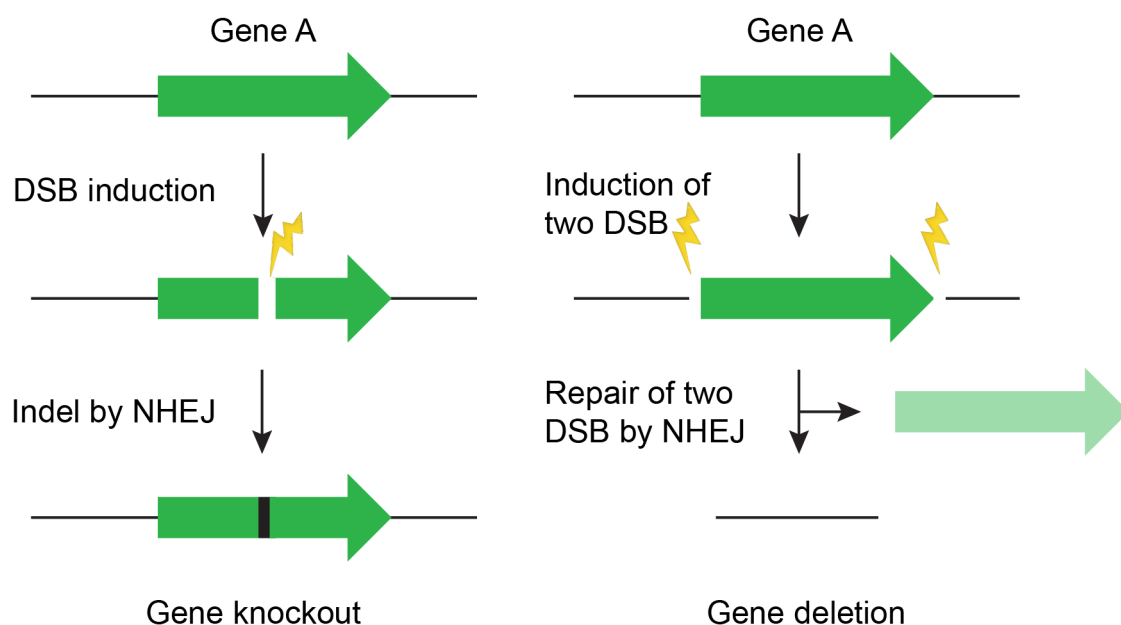


**Fig. 1.7 Homologous recombination can be divided in three phases: Presynapsis, Synapsis and Postsynapsis. Here one of the three possible sub-pathways, synthesis-dependent strand annealing is shown. This is the predominant pathway in somatic cells to repair DSBs. During presynapsis DSBs are recognized (A) and 3'-single strand overhangs are processed (B). In Synapsis after searching for homology strand invasion starts resulting in the D-loop (C). The invaded strand is extended by a DNA polymerase (D) and re-anneals to the second single strand of the DSB by complementary base pairing (E). Finally DNA synthesis fills up the remaining single-stranded gaps (F). Figure adopted from Carroll 2011.**

### 1.4.3 Genome engineering

DNA DSBs are one of the most severe injuries to DNA. They are repaired either by the NHEJ or the HR pathway. Both pathways can be exploited for genome engineering. ZFNs and TALENs were used to induce DSBs to trigger these repair pathways in a desired

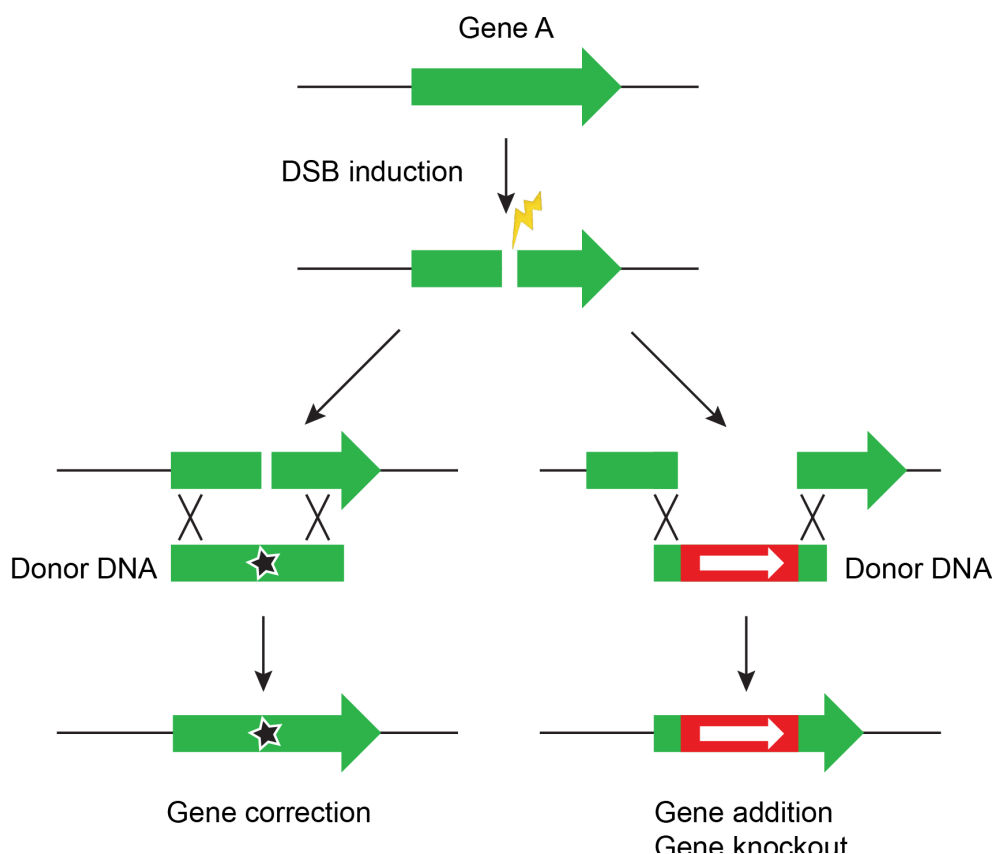
manner (GAJ *et al.* 2013). To achieve an induced DSB ZFNs or TALENs have to be transiently expressed in cells by introducing their encoding genes on expression vectors. They are transiently transfected into cells by chemical transfection, electroporation and integrase-deficient lentiviral-delivery (LOMBARDO *et al.* 2007; MAEDER *et al.* 2008; KIM *et al.* 2009). In the absence of a suitable homologous template the DSBs will be repaired by the cellular NHEJ pathway. A repair via NHEJ is error prone and will lead to a insertion or deletion of nucleotides (LIEBER and KARANJAWALA 2004). A DSB within the coding sequence of a gene will likely result in a frameshift mutation and thus a truncated and/or nonfunctional protein (Fig. 1.8 left; PORTEUS and CARROLL 2005). Both ZFNs and TALENs were used to knockout the HIV co-receptor CCR5 generating a HIV resistant genotype *de novo* in primary CD4+ T-cells and HEK293T cells, respectively (PEREZ *et al.* 2008; MUSSOLINO *et al.* 2011). If two DSBs are introduced on the same chromosome the sequence in between will be lost upon NHEJ repair resulting in gene deletion (Fig 1.8 right). Two pairs of designed ZFNs that lack cross-reactivity and recognize two adjacent sites in the human HOXB13 locus generated a targeted chromosomal deletion within exon 1 of the HOXB13 gene upon co-expression (SOLLU *et al.* 2010).



**Fig 1.8 Gene knockout or deleterious strategies with ZFNs or TALENs utilizes DSB repair by NHEJ. Directed DSBs are induced by ZFNs or TALENs triggering the NHEJ in the absence of a DNA repair template. A DSB induced within a gene will likely introduce a frameshift mutation (gene knockout). Two DSBs likely induce the loss of the sequence in between them deleting the gene. Figure modified from Kim and Kim 2011.**

Genomic engineering also utilizes the HR repair pathway in conjunction with templates supporting homologous DNA repair. Without a DNA lesion HR is very inefficient but a

DSB dramatically increases HR at the site of DNA damage. Delivering the DNA template together with a TALEN or ZFN pair for the induction of a specific DSB at the site of interest enhances the frequency of HR 100 to 1,000-fold in various different cell lines (JOHNSON and JASIN 2001). The outcome of the HR repair is determined by the architecture of donor DNA. The donor DNA can be designed to correct a mutation or insert a specific mutation into the genomic locus (Fig. 1.9 left). For example, gene correction was applied in human T cells to correct a mutation in the IL2R gene linked with the X-linked severe combined immune deficiency (SCID) using ZFNs (URNOV *et al.* 2005). The donor DNA can also be constructed to add a missing part of or even substitute an entire gene locus by flanking a gene of interest with sequences homologous to the DSB site (Fig. 1.9 right). TALENs were used to insert the *egfp* coding sequence in the tyrosine hydroxylase locus in zebrafish (ZU *et al.* 2013). This strategy can also be applied to achieve a gene knockout by inserting a DNA fragment into the coding region of a gene.



**Fig. 1.9 Exploiting the HR pathway for genomic engineering.** TALENs or ZFNs are used together with donor DNA, which serves as a repair template. The design of the donor DNA determines the outcome. Mutations can be corrected or inserted (left). Donor DNA with a desired sequence flanked by homologous arms can be used to insert a new gene or disrupt an endogenous gene (right). Figure modified from Kim and Kim 2011.

## 1.5 Murine gammaherpesvirus 68 (MHV-68)

In the framework of this thesis TALENs were used to establish an anti-viral system. The murine gammaherpesvirus 68 served as a model organism to test the novel system. Herpesviral genomes consist of double-stranded linear DNA with a length between 120 to 230 kilo bases, which is rapidly circularized after entry into the nucleus (POFFENBERGER and ROIZMAN 1985). The genome is contained in an icosahedral capsid surrounded by the tegument and then by a lipid envelope containing membrane-associated proteins (ROIZMAN and PELLET 2007). Upon primary infection herpesviruses establish a reservoir of latently infected cells and persist lifelong in their host evading the immune system. Occasionally, the viral lytic replication phase is reactivated and progeny virus is produced to maintain the viral reservoir in the host and to spread between hosts (YOUNG *et al.* 2007).

In humans the gamma-1 herpesvirus, Epstein-Barr virus (EBV) was the first virus linked to cancer (EPSTEIN *et al.* 1964) and is associated with lymphomas and nasopharyngeal carcinoma. For this reason EBV is classified as a type 1 cancerogenic agent by the International Agency for Research on Cancer (IARC). The gamma-2 herpesvirus Kaposi's sarcoma-associated herpesvirus (KSHV) causes the Kaposi's sarcoma and is regarded as the aetiologic agent of primary effusion lymphomas and is also classified as cancerogenic to humans (TAYLOR and BLACKBOURN 2011). Due to the restricted host range of these viruses *in vivo* studies are limited to clinical investigation. The murine gammaherpesvirus 68 (MHV-68) is a natural pathogen of wild murine rodents and is considered as a model for herpesvirus infections (ADLER *et al.* 2000). The MHV-68 genome has a length of about 120 kb consisting of 80 open reading frames (ORF). 63 of these ORFs have homologs in the KSHV genome and a majority is also present in the EBV genome (VIRGIN *et al.* 1997). This makes MHV-68 a model for gammaherpesvirus infections.

The cloning of a MHV-68 genome as bacterial artificial chromosome (BAC) has been described (ADLER *et al.* 2000). This approach allows the maintenance of the viral genome as a BAC in *E.coli* and a straightforward mutagenesis of the virus genome using the bacterial recombination machinery. Progeny virus is reconstituted by transfecting the

BAC plasmid into eukaryotic cells. In my work I made use of a virus termed MHV-68-gfp. It contains the BAC cassette and the coding sequence of the *gfp* gene both flanked by loxP sites, which were inserted into the viral genome. After passaging the MHV-68-gfp virus in fibroblasts expressing the Cre recombinase the inserted sequence is removed. This generates a virus, MHV-68, which is similar to the parental virus except that it contains a single Cre site, but lacks the BAC cassette and the *gfp* sequence. This MHV-68 revertant shows biological properties comparable to wild type MHV-68 (ADLER *et al.* 2001). Together both viruses, MHV-68-gfp and MHV-68, provide a model system, which allows assessing a possible TALEN-mediated effect on viral *de novo* synthesis in the framework of this thesis.

## 1.6 Aims

Deciphering the code that underlies the DNA binding of TAL effector proteins made the construction of customized DNA binding domains accessible to many researchers (BOCH *et al.* 2009; MOSCOU and BOGDANOVE 2009). The combination of the DNA binding domain of TAL effectors with the nuclease domain of the FokI restriction enzyme provided the basis for the easy construction of artificial nucleases with desired target sequences. Up to now the application of TALENs concentrated on genome engineering of cells and organisms. The aim of this thesis was to assess whether TALENs could be developed to act as antiviral factors.

The sequence of double-stranded DNA viruses has no or only very little homology to the genome of their host cells, which offers a unique opportunity to engineer TALENs targeting viral sequences, only. A TALEN pair specific for a certain DNA virus has to be expressed in cells, which are subsequently infected with this virus containing the target sequence of the TALENs. We postulated that upon entry of the viral genome into the nucleus of the infected cell a TALEN-mediated cleavage of the viral genome occurs. A DSB within the genome will most likely prevent the replication of the viral genomic DNA blocking viral *de novo* synthesis.

Together with Orlando de Lange (Laboratory Thomas Lahaye, LMU), I constructed two TALEN pairs specific for the GFP encoding sequence. Both pairs were compared for cleavage activity and the pair showing the higher activity was chosen for my further experiments. The herpes virus MHV-68-gfp served as a model for my infection experiments. To clarify if TALENs do indeed cut viral DNA as anticipated, I assessed the knockout of a virally encoded *gfp* gene at the level of DNA and protein expression. Whether an induced DNA double-strand break within the viral genome blocks viral *de novo* synthesis was the key question of this thesis. A second focus was on possible TALEN-mediated genotoxic effects in cells, which express the TALENs. Genotoxicity might result from TALEN-mediated cleavage of genomic DNA at off-target sequences. Little is known today about the stringency and specificity with which TALEs bind to their cognate sequence motifs as TALENs are primarily used in genomic engineering. In this particular approach only a limited dose of nucleases are expressed for a short period of time to promote two cleavage events, one in each allele, only. In order to generate cells, which are resistant to a certain virus, the antiviral nucleases have to be constitutively expressed. A constitutive expression bears the risk of unspecific cleavage events within the cellular genome resulting in genotoxic stress and genomic instability. This aspect is a major aspect of my work and experiments were performed to elucidate possible detrimental effects by TALENs on the cells.

When I started my project, very little was known about the biochemical characteristics of TALEs and zinc fingers. I therefore invested into analyzing the binding and dimerization characteristics of single ZFNs and a ZFN pair as well as the binding characteristics of a natural TALE and a synthetic TALE protein.

## 2. Materials

### 2.1 Oligonucleotides

All oligonucleotides used in this work were synthesized by the company Metabion (Munich). A list with their sequences is located in the appendix.

### 2.2 Plasmids

All plasmids listed below were generated, unless otherwise noted, in the context of this work.

plasmid	description
p3925	Flag and tandem Strep-tag expression plasmid; CMV promoter driven (GLOECKNER <i>et al.</i> 2007)
p4604	ZFN1 designed by Toni Cathomen (Universitätsklinikum Freiburg); codon-optimized at Geneart with point mutation D450A in FokI sequence, which eliminates cleavage activity (ZFΔN1)
p4605	ZFN2 designed by Toni Cathomen (Universitätsklinikum Freiburg); codon-optimized at Geneart with point mutation D450A in FokI sequence, which eliminates cleavage activity (ZFΔN2)
p4606	ZFΔN1 in 3925
p4607	ZFΔN2 in 3925
p4608	ZFΔN1-RR
p4609	ZFΔN2-DD
p4669	ZFΔN1-RR in 3925
p4670	ZFΔN2-DD in 3925
p4700	oriP wild type plasmid with FKBP12:DD-EBNA1:HMGA1a encoding mCherry
p4720	HAX3 TALE in 3925 fused to VP16 domain (Martin Bergbauer)
p4760	TALEN backbone designed by Toni Cathomen (Universitätsklinikum Freiburg) with terminal half repeat NI
p4761	TALEN backbone designed by Toni Cathomen (Universitätsklinikum Freiburg) with terminal half repeat NG
p4764	p3925 but MCS has been modified to be suitable for TALEN insertion



p4812	pRTR oriP expression vector with bidirectional doxycycline inducible promoter driving NGFR-IRES-eGFP and Luciferase (BORNKAMM <i>et al.</i> 2005)
p4839	TALEN 5 DNA binding domain module A5 (Thomas Lahaye – LMU)
p4840	TALEN 5 DNA binding domain module 5B (Thomas Lahaye – LMU)
p4841	TALEN 5 DNA binding domain module BC (Thomas Lahaye – LMU)
p4842	TALEN 6 DNA binding domain module A5 (Thomas Lahaye – LMU)
p4843	TALEN 6 DNA binding domain module 5B (Thomas Lahaye – LMU)
p4844	TALEN 6 DNA binding domain module BC (Thomas Lahaye – LMU)
p4845	TALEN 7 DNA binding domain module A5 (Thomas Lahaye – LMU)
p4846	TALEN 7 DNA binding domain module 5B (Thomas Lahaye – LMU)
p4847	TALEN 7 DNA binding domain module BC (Thomas Lahaye – LMU)
p4848	TALEN 8 DNA binding domain module A5 (Thomas Lahaye – LMU)
p4849	TALEN 8 DNA binding domain module 5B (Thomas Lahaye – LMU)
p4850	TALEN 8 DNA binding domain module BC (Thomas Lahaye – LMU)
p4852	TALEN 5 binding sites in p4760
p4853	TALEN 7 binding sites in p4760
p4854	TALE 5 (without FokI) in p4764
p4865	TALEN 6 binding sites in p4760
p4866	TALEN 8 binding sites in 4761
p4867	TALEN 7 in p4764
p4879	TALEN 5 in p4764
p4880	TALEN 6 in p4764
p4881	TALEN 8 in p4764
p4894	ZFN1 expression plasmid with activated FokI nuclease domain (originates from p4606)
p4895	ZFN2 expression plasmid with activated FokI nuclease domain (originates from p4607)
p5252	pRTR vector (based on p4812) expresses TALEN 7-IRES-cerulean and TALEN 8 upon induction with dox
p5257	pRTR vector (based on p4812) expresses mCherry and luciferase upon induction with dox
p5387	eGFP expression vector with HAX3 binding site within eGFP sequence
p5388	TALEN backbone p4760 with R487D mutation in FokI sequence
p5389	TALEN backbone p4760 with E490K mutation in FokI sequence
p5390	TALEN backbone p4760 with Q486E mutation in FokI sequence
p5407	HAX3 expression plasmid based on p4720 but VP16 domain has been removed and C-term is truncated like in p4854
p5408	TALEN backbone p4760 with E490K and I538K mutation in FokI sequence
p5409	TALEN backbone p4760 with Q486E and I499L mutation in FokI sequence
p5420	TALEN backbone p4760 with D483R mutation in FokI sequence
p5434	TALEN 7-RR in p4764
p5435	TALEN 8-DD in p4764
p5436	TALEN 7-KK in p4764
p5437	TALEN 8-EL in p4764
p5471	TALEN 7-ELD in p4764
p5472	TALEN 8-KKR in p4764
p5473	TALEN 7-KKR in p4764

p5474	TALEN 8-ELD in p4764
p5480	pRTR vector (based on p4812) expresses TALEN 7-KKR-IRES-cerulean and TALEN 8-ELD upon induction with dox
p5567	pRTR vector (based on p4812) expresses IRES-Cerulean upon induction with dox
p5568	pRTR vector (based on p4812) expresses TALEN 7-KKR-IRES-NGFR and TALEN 8-ELD upon induction with dox

## 2.3 Antibodies

All antibodies that were used in this work are listed below.

specificity	species	source	application	comment
anti-Flag	mouse	Stratagene (#200472)	WB 1:1000	-
anti-HA	mouse	E. Kremmer	WB 1:10	12CA5
anti-tubulin	mouse	Santa Cruz (#sc 23948)	WB 1:2500	-
anti-mouse IgG	horse	Cell Signaling (#7076S)	WB 1:5000	HRP conjugated
anti-mouse IgG	goat	Caltag (#30005)	FACS 1:200	APC conjugated
anti-P-histone H2A.X	rabbit	Cell Signaling (#9719S)	FACS 1:50	Alexa 488 conjugated
anti-CD2	mouse	E. Kremmer	MACS 400µl	OX34-1
anti-human NGFR	mouse	BioLegend (#345108)	FACS 1:20	APC conjugated

## 2.4 Bacterial strains

For cloning and plasmid DNA preparation the bacterial strain DH5-alpha was used. In case plasmid DNA without Dam and Dcm methylation was required the dam-/dcm- strain GM2163 was used.

name	genotype
DH5-alpha	F-, lacI-, recA1, andA1, hsdR17, Δ(lacZYA- argF), U169, F80d lacZ ΔM15, supE44, thi-1, gyrA96, relA1 (HANAHAN 1985)
GM2163	F-, ara-14, leuB6, tonA31, lacY1, tsx-78, supE44, galK2, galT22, hisG4, rpsL136, xyl-5, mt1-1, thi-1, dam-13, Tn9, dcm6, hsdR2, mcrA-, mcrB-

## 2.5 Eukaryotic cell lines

Cell lines, which were used in this work, stem from the collection of the Department of Gene Vectors or were established in the context of this work (\*). In case of established cell lines the plasmid number with which the cells were stably transfected is indicated.

name	description
HEK293	Human embryonic kidney cells transformed after transfection with DNA from human Adenovirus Type 5 (GRAHAM <i>et al.</i> 1977)
K562	Immortalized human cell line derived from a patient with chronic myeloid leukemia (LOZZIO and LOZZIO 1975)
BHK21	Fibroblast cell line derived from a baby hamster kidney (STOKER and MACPHERSON 1964)
U-2 OS (GFP <sup>+</sup> )	Human osteosarcoma cell line derived from a moderately differentiated sarcoma from the tibia (PONTEN and SAKSELA 1967), which harbors a genomically integrated copy of the GFP sequence (Toni Cathomen - Universitätsklinikum Freiburg).
HEK293/5252*	HEK293 cells stably transfected with p5252 expressing TALEN7, TALEN8 and Cerulean
HEK293/5257*	HEK293 cells stably transfected with p5257 expressing mCherry and luciferase
HEK293/5480*	HEK293 cells stably transfected with p5480 expressing TALEN7-KKR, TALEN8-ELD and Cerulean

## 2.6 Cell culture media and additives

### 2.6.1 Media for the cultivation of prokaryotic cells

Bacteria were either cultivated in liquid culture with LB-medium for large scale production or on solid agarose plates if single clones were cultured.

description	ingredients	distributor
LB-medium	1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract	Invitrogen GmbH
select agar	LB-medium with 1.5% (w/v) select agar	Invitrogen GmbH

### 2.6.2 Media for the cultivation of eukaryotic cells

description	application	distributor
RPMI-1640	cell culture medium for the cultivation of HEK293 cells and stably transfected cell lines	Invitrogen GmbH
DMEM	cell culture medium for the cultivation of U-2 OS cells	Invitrogen GmbH
Glasgow MEM (BHK 21)	cell culture medium for the cultivation of BHK21 cells	Pan Biotech GmbH
Opti-MEM I	reduced serum medium used to set up transfections	Invitrogen GmbH
fetal calf serum (FCS)	nutritive additive for RPMI-1640, DMEM and Glasgow MEM (BHK 21)	Invitrogen GmbH
sodium pyruvate	antioxidant agent for RPMI-1640	Invitrogen GmbH
sodium selenite	antioxidant agent for RPMI-1640	Sigma Aldrich
alpha-thioglycerole and bathocupproine-disulfonic acid	antioxidant agents for RPMI-1640	Sigma Aldrich
tryptose phosphate broth (TBP)	nutritive additive for Glasgow MEM (BHK 21)	Sigma Aldrich
l-glutamine	nutritive additive for Glasgow MEM (BHK 21)	Sigma Aldrich
penicillin/streptomycin	antibiotics for RPMI-1640, DMEM and Glasgow MEM (BHK 21)	Invitrogen GmbH
puromycin	antibiotic used for selection of stably transfected cell lines	

## 2.7 Chemicals and Enzymes

distributor	description
Affymetrix, Santa Clara, USA	T4 DNA Ligase
Applichem, Darmstadt	RNAse A, Lysozyme
Bio Rad, Munich	Bradford
Biontex, Martinsried	Metafectene
Carl Roth GmbH, Karlsruhe	phenol, chloroform, bovine serum albumin fraction V (BSA), Rotiphorese Gel30 (acrylamide), powdered milk
Difco Laboratories, Detroit, USA	bacto-agar, yeast extract, trypton
Fermentas, St. Leon-Rot	DNA markers, loading dye for gel electrophoresis, restriction enzymes, T4 Polynucleotide Kinase
Finnzymes, Vantaa, Finland	Phusion polymerase
GE Healthcare, Munich	Hybond-ECL membrane

Hartmann Analytik, Braunschweig IBA GmbH, Göttingen	[ $\gamma$ -P <sup>32</sup> ] dATP Strep-Tactin Sepharose, desthiobiotin elution buffer (10x), wash buffer (10x)
Invitrogen, Karlsruhe Merck-Eurolab GmbH, Darmstadt	agarose, trypsin-EDTA ethanol, ethidium bromide, ethylene diamine tetra-acetic acid (EDTA), glycine, glycerol, isopropanol, methanol, magnesium chloride (MgCl <sub>2</sub> ), sodium chloride (NaCl), sodium hydroxide (NaOH), sodium dodecyl sulfate (SDS), hydrochloric acid (HCl), Triton X-100, acetic acid, N,N,N',N'- Tetramethylethylenediamin (Temed)
Miltenyi Biotec, Bergisch Gladbach New England Biolabs, Frankfurt	Anti-mouse IgG Micro Beads restriction enzymes, Klenow fragment, T4 DNA polymerase
Roche Diagnostics GmbH, Mannheim	alkaline phosphatase, Complete Mini Protease Inhibitor, desoxy nucleotides, PhosStop Phosphatase Inhibitor, LightCycler 480 SYBR Green I Master, FugeneHD, poly(dI/dC)
SERVA Electrophoresis GmbH, Heidelberg Sigma-Aldrich, Munich	ammonium persulfate (APS) ampicillin, beta-mercaptoethanol, bromphenol blue, dimethyl sulfoxide (DMSO), dithiothreitol (DTT), doxycycline, puromycin, Ponceau S (Fluka ATX), carboxymethylcellulose (400-800 cps), polyethylenimine (PEI), Igepal CA-630 (NP40), Monolaurat
Thermo Scientific, Waltham, USA	PageRuler Plus Prestained Protein Ladder

## 2.8 Buffers and solutions

buffer/solution	ingredients	application
agarose loading dye	1 mg/ml bromphenol blue, 1 mg/ml xylene cyanol, 50 % sucrose	commonly used buffers and solutions
ampicillin 1000x	100 mg/ml ampicillin in 50 % ethanol, sterile filtered	
kanamycin 1000x	30 mg/ml kanamycin in ddH <sub>2</sub> O, sterile filtered	
PBS, pH 7.3	10 mM Na <sub>2</sub> HPO <sub>4</sub> , 1.8 mM KH <sub>2</sub> PO <sub>4</sub> , 140 mM NaCl, 2.7 mM KCl	
TAE	40 mM Tris-acetate,	

	1mM EDTA	
TBE	89 mM Tris, 89 mM boric acid, 2 mM EDTA	
TE, pH 8.0	10 mM Tris-HCl, 1 mM EDTA	
M-STET	0.5 % Triton-X-100, TrisHCl pH 8.0, 5 % sucrose, 25 mM EDTA	
running buffer 10x	30 g Tris/Base, 144 g glycine, 100 mL 10 % SDS, add ddH <sub>2</sub> O to 1 l	Western blot
blotting buffer 1x	3 g Tris/Base, 14.4 g glycine, 200 ml methanol, add ddH <sub>2</sub> O to 1 l	
washing buffer (PBS-T) blocking solution	0.1 % Tween-20 in PBS 5 g milk powder in 100 ml PBS-T	
Laemmli buffer 5x	2.5 ml 2 M Tris-HCl, pH 6.8, 1 g SDS, 5 ml 50 % glycerol, 0.78 g DTT, 0.2 % bromphenol blue, add H <sub>2</sub> O to 10 ml	
acrylamide stacking gel 5%	1.34 ml 30 % acrylamide, 1.5 ml 2 M Tris-HCl, pH 6.8, 80 µl 10 % SDS, 7.2 ml H <sub>2</sub> O, 100 µl 10 % APS, 10 µl TEMED	
acrylamide separation gel 10%	3.3 ml 30 % acrylamide, 1.66 ml 2 M Tris-HCl, pH 8.9, 67 µl 0.5 M EDTA, 100 µl 10 % SDS, 4.9 ml H <sub>2</sub> O, 70 µl 10 % APS, 10 µl TEMED	
binding buffer 10x	200 mM HEPES, pH 7.9, 10 mM DTT, 20 mM MgCl <sub>2</sub> , 10 % glycerol, 0.1 mg/ml BSA, + protease inhibitor	EMSA
dilution buffer	10 % 10x binding buffer, 67 % washing buffer from protein purification, 23 % ddH <sub>2</sub> O	
native 10 % acrylamide gel	6.7 ml 30 % acrylamide, 1 ml 10x TBE, 160 µl APS, 20 µl TEMED, 12.2 ml ddH <sub>2</sub> O	
fixing solution	10 % methanol, 20 %	

	acetic acid in ddH <sub>2</sub> O	
washing buffer	100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA	Protein purification
elution buffer	+ protease inhibitor 100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, +protease inhibitor	
RIPA	50 mM Tris-HCl, pH 8.0, 150mM NaCl, 0.1 % SDS, 1 % NP-40, 0.5 % DOC + protease inhibitor	Cell lysis
cell suspension/washing buffer	2 % FCS, 2 mM EDTA in PBS	Flow cytometry / MACS cell separation
washing buffer	0.1 % NaN <sub>3</sub> (10% solution), 5 % FCS in PBS	Cell permeabilization

## 2.9 Commercial kits

distributor	description
GE Healthcare, Munich	Illustra MicroSpin G50 columns
Genomed, Bad Oeynhausen	Jet Star Plasmid Purification Kit
Invitrogen GmbH, Karlsruhe	Fix&Perm Cell Permeabilization Kit
Macherey-Nagel, Düren	NucleoSpin Gel and PCR Clean-up
Promega, Mannheim	TNT Quick Coupled Transcription/Translation System
Qiagen, Hilden	Endofree Plasmid Maxi Kit, QIAprep Spin Miniprep Kit, QIAamp DNA Mini Kit
Roche Diagnostics GmbH, Erlangen	Rapid DNA Dephos and Ligation Kit

## 2.10 Devices and consumables

distributor	description
ACEA, San Diego, USA	e-plate 16
AGFA, Cologne	CP1000 developer machine, Medical X-Ray Screen
Beckman, Heidelberg	centrifuge Avanti J-25 and Avanti-26XP
Becton Dickinson GmbH, Heidelberg	LSRFortessa, FACS-Aria III, cell culture flasks
Bio-Rad, Munich	Mini Protean II Tetra Gelelectrophoresis Unit
Brand, Wertheim	glass pipettes 5 ml, 10 ml, 20 ml and 25 ml
Branson, Danbury, USA	Branson Sonifier S-250D
Eppendorf, Hamburg	BioPhotometer, PCR machine
	Mastercycler Personal, reaction tubes, desk centrifuge 5415, thermomixer comfort
Eurofins MWG Operon, Ebersberg	DNA sequencing
Hartenstein, Würzburg	cover slips, glass slides, forceps, whatman paper, Neubauer cell chamber
Hewlett Packard, USA	HP Scan-Jet G4050
G. Heinemann, Schwäbisch Gmünd	sonifier nozzle 5 mm for Branson sonifier
FujiFilm, Kleve	Phosphorimager FLA 5100, film cassettes, intensifier screen
Invitrogen GmbH, Karlsruhe	Qubit 2.0 Fluorometer, Neon Transfection System
GE Healthcare, Munich	High performance autoradiography film
Gilson Inc., Middleton, USA	2 µl, 20 µl, 200 µl and 1000 µl pipettes
Greiner Bio One, Solingen	plastic consumables for cell culture and laboratory work
Metabion, Martinsried	DNA oligonucleotides and primers
Milipore, Schwalbach	water deionizing device Milli-RO 60 PLUS
Miltenyi Biotec, Bergisch Gladbach	MACS separation columns LS
Montreal Biotec Inc., Montreal, Canada	vacuum centrifuge univapo H150
Nunc GmbH, Wiesbaden	cell culture dishes, 6-/12-/24-/48- and 96-well plates, cryotubes 1.5 ml
Peqlab, Erlangen	agarose gel electrophoresis chambers, NanoDrop ND-1000, reaction tubes
Roche Diagnostics GmbH, Erlangen	Light Cycler 480 Real Time PCR System, Xcelligence RTCA DP
Schleicher & Schüll, Dassel	0.8 µm and 1.2 µm filter
Schott, Mainz	glass ware
Sequisservice GmbH, Vaterstetten	DNA sequencing
Stratagene, Heidelberg	Robocycler Gradient 96
Thermo Scientific, USA	Haeraeus Multifuge 3L-R, orbital shakers, incubators
Vilber Lourmat, Marne-la-Vallée, France	Quantum ST4
Zeiss, Göttingen	Axiovert 40C, Axiovert 200M



## 2.11 Software

distributor	description
Accelrys, Cambridge, UK	MacVector 12.7
Adobe Systems Inc., San Jose, USA	Illustrator CS5, Photoshop CS5
Graphpad Software, La Jolla, USA	Prism 6
Microsoft, Redmond, USA	Office 2011
Raytest Isotopenmessgeräte GmbH, Straubenhardt	AIDA Alias Software
Thomson Reuters, New York, USA	EndNote X6
TreeStar Inc., Ashland, USA	FlowJo 9.5
Vilber Lourmat, Marne-la-Vallée, France	Quantum ST4 software

## **3. Methods**

### **3.1 Bacterial culture**

#### **3.1.1 Growth and storage of bacterial cultures**

*Escherichia coli* (E.coli) was cultured in LB medium suspension at 37 °C and 200 rpm in orbital shakers. Cultures were inoculated using a sterilized inoculating loop. Single cell clones were obtained by streaking the bacteria on LB agar using an inoculating loop or a glass spreader and incubating the agar plates at 37 °C for 12-16 h. Antibiotics (100 µg/ml ampicillin or 30 µg/ml kanamycin) were added to the LB medium after autoclaving to select for bacteria harboring plasmid DNA with resistance genes. For long-term storage of bacteria stocks, over night bacteria were centrifuged, resuspended (50 % LB medium, 25% PBS and 25% glycerol) and stored at -80 °C.

#### **3.1.2 Transformation of bacteria**

Chemical competent E.coli cells were obtained according to the protocol of Hanahan et al. (HANAHAN 1985). For long-term storage 200 µl aliquots were made, frozen immediately in liquid nitrogen and transferred to a -80 °C freezer.

For a single transformation reaction bacteria were thawed on ice, 50-100 µl were incubated with (i) 2 µl (equivalent to 5-10 ng) of a ligation reaction or (ii) 1 ng of plasmid DNA on ice for 20 minutes. The reaction was subsequently incubated on 42 °C

for 90 seconds, transferred back on ice and 900 µl of LB medium was added. The mixture was incubated at 37 °C to allow expression of resistance genes. After 30 minutes bacteria were centrifuged at 3000 g for 30 seconds, resuspended in 100 µl and plated on LB agar containing the adequate antibiotic. The LB agar plates were incubated at 37 °C for 12-16 h.

## **3.2 Eukaryotic cell culture**

### **3.2.1 Cell culture conditions**

Cell culture was done in lamina hoods using sterile pipettes, solutions and culture dishes. Cells were cultivated in incubators at 37 °C in an atmosphere of 5 % CO<sub>2</sub> and 95 % humidity. Centrifugation steps were carried out at 300 g for 5 minutes. Cells were washed with sterile PBS. Cell number was determined using a Neubauer cell chamber.

#### HEK293 cells / stably transfected cell lines

Adherent HEK293 cells were cultured in RPMI1640 supplemented with 10 % FCS, 1 mM sodium pyruvate, 100 nM sodium selenite, 20 µM bathocupproine-disulfonic acid, 0.433 % α-thioglycerole, 100 µg/ml streptomycin and 100 Units/ml penicillin. When cells reached a confluency of 70-80 % they were washed with 10 ml PBS and subsequently incubated with 2 ml of 0.05 % Trypsin/0.02 % EDTA. The detached cells were resuspended in 10 ml medium and 1 ml was transferred to a new culture vessel. Stable HEK293 cell lines were cultivated with 500 ng/ml puromycin.

#### U-2 OS cells

U-2 OS cells were cultured in DMEM supplemented with 10 % FCS, 100 µg/ml streptomycin and 100 Units/ml penicillin. When cells reached a confluency of 70-80 % they were passaged in a 1:15 ratio like described above.

### BHK21 cells

BHK21 cells were cultivated in Glasgow MEM (BHK21) supplemented with 5 % FCS, 5 % TPB, 1 % l-glutamine, 100 µg/ml streptomycin and 100 Units/ml penicillin. When cells reached a confluency of 70-80 % they were passaged in a 1:20 ratio like described above. If cells were used in plaque assays with titrated viral supernatants the normal BHK21 culture medium was exchanged by a carboxymethylcellulose overlay medium after cells were infected with MHV-68. To prepare the overlay medium 320 ml of ddH<sub>2</sub>O with 3 g of carboxymethylcellulose and a magnetic stir bar was autoclaved. After autoclaving the solution was stirred until carboxymethylcellulose was completely dissolved. Before use the solution was supplemented with 40 ml 10x BHK21 medium, 20 ml FCS, 20 ml TPB, 4 ml penicillin/streptomycin, 4 ml L-glutamine and 13.1 ml 1M NaHCO<sub>3</sub>.

### K562 cells

K562 cells were cultured in suspension in RPMI1640 supplemented with 10 % FCS, 1 mM sodium pyruvate, 100 nM sodium selenite, 20 µM bathocupproine-disulfonic acid, 0.433 % α-thioglycerole, 100 µg/ml streptomycin and 100 Units/ml penicillin. They were passaged every 3-4 days in a 1:10 ratio.

## **3.2.2 Storage of eukaryotic cells**

For long term storage cells were trypsinized, centrifuged and resuspended in a mixture of 90 % FCS and 10 % DMSO at a concentration of  $1 \times 10^7$  cells/ml. The suspension was transferred to 2 ml CryoTubes (Nunc) and gently frozen at -80 °C using a freezing container (Nunc) filled with isopropyl alcohol, resulting in a cooling rate of 1 °C/min. After one week one vial was thawed in order to test cell viability. The other vials were transferred to liquid nitrogen tanks and stored in the gas phase.

To thaw cells, vials were put in a water bath at 37 °C until they were completely thawed and subsequently transferred to 50 ml of fresh medium. Cells were centrifuged, resuspended in medium and plated in a culture vessel. If antibiotics had to be added this was done two days after thawing the cells.

### 3.2.3 Transfection of HEK293 cells

HEK293 cells were transfected with nucleic acids using Metafectene (Biontex) or PEI (Sigma-Aldrich) depending on the cells.

#### 6-Well plate (Metafectene)

$5 \times 10^5$  cells were seeded per well. The next day the following mixtures were set up:

100  $\mu$ l Opti-MEM I + 1  $\mu$ g DNA

100  $\mu$ l Opti-MEM I + 6  $\mu$ l Metafectene

Both solutions were mixed, incubated for 15 minutes at room temperature and then added to the cells. The mixture was left on the cells over night.

#### 15 cm cell culture dish (PEI)

$2 \times 10^7$  cells were seeded per dish. The next day culture medium was replaced by 10 ml Opti-MEM I and the following mixtures were set up:

3 ml Opti-MEM I + 15  $\mu$ g DNA

3 ml Opti-MEM I + 90  $\mu$ l PEI (1 mg/ml)

Both solutions were mixed, incubated for 15 minutes at room temperature and then added to the cells. After 5 hours the transfection mixture was replaced by 20 ml of culture medium.

### 3.2.4 Electroporation of U-2 OS / K562 cells

U-2 OS and K562 cells were electroporated using the Neon transfection device (Invitrogen).

### U-2 OS cells

6-Well plates were prepared with 2 ml medium per well. Cells were trypsinized, counted, washed in PBS and  $1 \times 10^6$  cells were resuspended in 100  $\mu$ l buffer R. 6  $\mu$ g of plasmid DNA was added to this solution. The electroporation was performed using the 100  $\mu$ l pipette and the following settings:

1200 V, 10 ms and 4 pulses

After electroporation the mixture was transferred to the prepared 6-well plates.

### K562 cells

6-Well plates were prepared with 2 ml medium per well. Cells were counted, washed in PBS and  $5 \times 10^5$  cells were resuspended in 100  $\mu$ l buffer R. 13  $\mu$ g of plasmid DNA was added to this solution. The electroporation was performed using the 100  $\mu$ l pipette and the following settings:

1350 V, 10 ms and 4 pulses

After electroporation the mixture was transferred to the prepared 6-well plates.

## **3.2.5 Generation of stable cell lines**

Two wells of a 6-well plate with HEK293 cells were transfected like described in chapter 3.2.3. One day after transfecting the cells they were trypsinized and transferred to one 15 cm cell culture dish with 20 ml medium and 500 ng/ml puromycin. Every 3-4 days medium was exchanged.

### 3.2.6 Staining of cells

Washing steps, antibody dilutions and cell resuspension for FACS analysis were all done in flow cytometry buffer, which is described above in 2.8.

#### ***3.2.6.1 Staining of cell surface markers***

HEK293 cells were stained for the surface markers CD2 and NGFR for FACS analysis according to the following protocols.

##### Staining for CD2

1x10<sup>6</sup> HEK293 cells were trypsinized, washed and resuspended in 50 µl OX34-1 (anti-CD2) antibody solution. After incubation for 20 minutes at 4 °C cells were washed two times and subsequently incubated in FACS buffer with 1:200 anti-mouse APC (Caltag M30005) for 20 minutes at 4 °C. After washing the cells were resuspended in 500 µl of FACS buffer and analyzed via flow cytometry. As a negative control CD2-negative HEK293 cells were used.

##### Staining for NGFR

1x10<sup>6</sup> HEK293 cells were trypsinized, washed, resuspended in 100 µl anti-human NGFR (1:20) antibody solution and incubated for 20 minutes at 4 °C. After two washing steps cells were resuspended in 500 µl FACS buffer and analyzed via flow cytometry. As a negative control NGFR-negative HEK293 cells were used.

#### ***3.2.6.2 Intracellular staining of phosphorylated histone H2A.X***

For intracellular staining the Fix & Perm Cell Permeabilization Kit (Invitrogen) was used. Washing steps were performed in washing buffer as described in 2.8. 100 µl of reagent A was mixed with 5x10<sup>5</sup> K562 cells in 100 µl PBS and incubated at room temperature for 15 minutes. After washing in 3 ml cells were resuspended in 100 µl reagent B including

2 µl of anti-P-histone H2A.X (Cell Signalling) and incubated for 20 minutes at room temperature. Then cells were washed once, resuspended in FACS buffer and analyzed via flow cytometry.

### **3.2.7 Flow cytometry**

Cells were resuspended in FACS buffer and analyzed with a LSRFortessa (Beckton Dickinson). In case different fluorochromes were measured at the same time, the different dyes were compensated against each other. Data was evaluated using the program FlowJo 9.5.

### **3.2.8 Generation of clonal cell lines**

Clonal cell lines were established by trypsinizing and seeding the cells on new 15 cm culture dishes in a 1:10, 1:100, 1:1000 and 1:10000 dilution. Cells were kept in culture until single colonies became visible to a size of 2 to 4 mm in diameter. Cells were washed once with PBS and covered with trypsin, which was aspirated immediately. Then small pieces of autoclaved whatman paper soaked in trypsin were put on distinct colonies using a forceps. After 5 minutes of incubation the snippets were transferred to 6-well plates with culture medium and kept in culture until cells started growing throughout the whole well.

### **3.2.9 Cell separation using MACS beads**

All washing steps were performed in washing buffer (see 2.8). One 15 cm plate of 70-80% confluent cells was used. Cells were sorted for CD2 expression using MACS beads by trypsinizing, washing and resuspending them in 400 µl anti-CD2 (OX34-1) antibody. After incubation for 20 minutes at 4 °C cells were washed two times in 3 ml washing buffer and resuspended in 160 µl washing buffer plus 40 µl of anti-mouse IgG micro



beads. Cells were incubated with the beads for 15 minutes at 4 °C, washed once and resuspended in 500 µl washing buffer. For the separation procedure LS columns were equilibrated in the magnetic separator with 3 ml buffer, cells were applied and the column was washed three times with 3 ml. For elution the column was taken out from the separator and 5 ml of buffer was applied and gently extruded using a plunger. The eluate was centrifuged and cells were plated in culture medium.

### **3.2.10 Cell sorting using BD FACSAria III**

HEK293 cells stably transfected with 5252/5480 were sorted for cerulean expression using FACSAria III. One day before sorting a 15 cm dish of cells, 70-80 % confluent, were induced with 100 ng/ml doxycycline. The next day cells were trypsinized, washed and taken up in FACS buffer. To obtain a single cell solution cells were filtered in FACS tubes with a cell strainer cap. Cells were sorted with the 130 µm nozzle and four-way purity settings.

### **3.2.11 Measurement of cell proliferation using the Xcelligence system**

Proliferation of HEK293 cells and different stably transfected cell lines in presence or absence of doxycycline was compared using the Xcelligence system. Cells were first MACS sorted (see 3.2.9) and counted. Then 5000 cells per well were plated after calibrating the wells with medium, only, and the measurement was initiated. Cells were induced 24 hours later by adding doxycycline to a final concentration of 100 ng/ml. In parallel, cells were plated and induced in 6-well plates in order to assess the induction rates via flow cytometry.

### **3.3 Virological techniques**

#### **3.3.1 Infection of HEK293 derived cell lines with MHV-68**

Cells were sorted using either MACS beads (5252/5257) or FACS Aria III (5252/5480) depending on the experiment. In case of MACS beads sorting the cells were first sorted, counted and then plated and induced. If cells were sorted using FACS Aria they were induced with doxycycline one day in advance and then sorted, counted and plated. In each case  $2 \times 10^5$  cells were used per well (12-well dish) in 1 ml of culture medium. After 24 hours cells were infected with a MOI of 0.1 by adding viral supernatant diluted in culture medium in a total volume of 100  $\mu$ l. After 90 minutes medium containing the virions was replaced by 2 ml of fresh culture medium without puromycin and with or without doxycycline, depending on the sample. Samples were collected after 24, 36, 48, and 60 hours by transferring the culture vessel to -80 °C.

#### **3.3.2 Generating viral supernatants**

Viral samples generated in 3.3.1 were thawed completely at room temperature and frozen again at -80 °C for at least one hour. Samples were thawed again and transferred to 15 ml falcon tubes and centrifuged for 10 minutes at 1100 g. After centrifugation 1 ml of each sample was transferred to 1.5 ml tubes and stored at -80 °C.

#### **3.3.3 Determination of viral titer**

Viral titer was assessed using BHK21 cells in 24-well plates. Per viral sample 6 wells were used to dilute the sample starting at 1:10 up to  $1:10^5$ . The serial dilution of the viral sample was done in a 200 $\mu$ l scale and added to the culture medium 24 hours after plating the cells for 90 minutes. Then the medium was aspirated and replaced by 2 ml of

overlay medium. After 5 days the medium was replaced by 300 µl crystal violet staining solution for 15 minutes. The cells were washed carefully with tap water and dried over night. The plaques were then counted under a phase contrast microscope.

### **3.4 Nucleic acid techniques**

General methods like phenol/chloroform extraction, DNA precipitation with ethanol or isopropanol, electrophoretic separation of DNA, enzymatic restriction hydrolysis and radioactive labelling of DNA were performed according to standard protocols (SAMBROOK 2001).

#### **3.4.1 DNA purification from E.coli**

##### Small-scale purification of plasmids for evaluation of cloning

Single colonies were streaked on LB agar using wooden autoclaved toothpicks. The agar plates were incubated overnight at 37 °C. Bacteria were harvested using toothpicks and resuspended in M-STET buffer, incubated at 95 °C for 2 minutes and centrifuged for 10 minutes at 16000 g.

##### Small-scale purification of plasmids

For small-scale purifications 7 ml of LB medium with bacteria in the stationary phase were centrifuged at 1000 g for 10 minutes. Plasmid DNA was purified using the Qiagen Spin Miniprep Kit.

##### Large-scale purification of plasmids

For larger amounts of DNA 400 ml of bacteria culture in the stationary phase was used. The purification was performed using either the Jet Star Plasmid Purification Kit (Genomed) for cloning or the EndoFree Plasmid Maxi Kit (Qiagen) for transfection of cells.

### **3.4.2 DNA purification from eukaryotic cells**

DNA from eukaryotic cells was purified using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's protocol.

### **3.4.3 DNA isolation from agarose gels**

DNA fragments were isolated using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer's protocol.

### **3.4.4 DNA isolation from SDS gels**

Annealed oligo nucleotides were made visible via an UV lamp. The fragment was excised and the gel piece was incubated in 100 µl TE in a 1.5 ml tube at room temperature on a thermomixer over night. On the next day the TE buffer was transferred to a new tube and the DNA concentration was measured.

### **3.4.5 Dephosphorylation and ligation**

DNA fragments were dephosphorylated and ligated using the Rapid DNA Dephos and Ligation Kit (Roche) according to the company's protocol.

### **3.4.6 Polymerase chain reaction (PCR)**

Polymerase chain reaction was used to exponentially amplify defined DNA fragments (MULLIS *et al.* 1986). PCR reactions were set up in a 50 µl scale (see below) and incubated in a PCR machine.

PCR reaction

component	amount
DNA	1-50 ng
5x Phusion buffer	10 $\mu$ l
forward primer	0.5 $\mu$ l (100 $\mu$ M)
reverse primer	0.5 $\mu$ l (100 $\mu$ M)
dNTPs	1 $\mu$ l (10 mM)
Phusion polymerase	0.5 $\mu$ l
ddH <sub>2</sub> O	ad 50 $\mu$ l

PCR program

temperature	time	cycles
98 °C	120 sec	1
98 °C	30 sec	30
50-60 °C (depending on primer pair)	30 sec	
72 °C	30 sec/kb	
72 °C	600 sec	1
6 °C	$\infty$	1

The amplified DNA was separated on a 1 % TBE agarose gel to control the size or extract the fragment.

### 3.4.7 Quantitative real time PCR (qPCR)

The ratio of MHV-68 coded genes ORF6 and GFP was compared using a LightCycler 480 (Roche). During DNA synthesis the fluorescent dye SYBR-Green I is incorporated into double-stranded DNA and therefore the amount of DNA can be quantified (HIGUCHI *et al.* 1993).

The relative quantification was performed according to the  $\Delta\Delta C_T$  Method (Applied Biosystems - Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR).

The qPCR reaction and qPCR program are listed below.

## qPCR reaction

component	amount
DNA	1 $\mu$ l
SYBR Green I Master (2x)	5 $\mu$ l
primer mix	0.5 $\mu$ l (5 pm each)
ddH <sub>2</sub> O	3.5 $\mu$ l

## qPCR program

program	target temperature (°C)	hold (s)	acquisition mode	RampRate (°C/s)	Cycles	Analysis mode
pre-incubation	95	600	none	4.4	1	none
amplification	95	1	none	4.4	45	quantification
	62	10	none	2.2		
	72	10	none	4.4		
	75	3	single	4.4		
melting curve	97	1	none	4.4	1	melting curve
	67	10	none	2.2		
	97	/	continuous	0.11		
cooling	37	15	none	2.2	1	none

### 3.4.8 Mutagenesis PCR

Single point mutations were introduced in plasmid DNA using the Fusion High-Fidelity DNA polymerase. Primers were designed such that their 5' ends are located close to each other, resulting in an amplification of the whole plasmid except the sequence between the primers. The sequence between the 5' ends of the primers was replaced by adding the desired sequence to the 5' end of one primer. The PCR mix is described in chapter 3.4.6 except that template amount was 10 ng. The PCR program differs only in cycle number, which was reduced to 25x.

### 3.4.9 Cut-ligation of TALEN

TALEN were assembled via the cut-ligation method, fusing together the binding modules (A5, 5B and BC) and the backbone (TALEN C-term/N-term and FokI domain). The cut-

ligation was performed as described below in a Mastercycler Personal PCR machine (Eppendorf). From each plasmid 40 fmol DNA was used. After the cut-ligation DNA was transformed in E.coli and clones were verified via restriction digest.

cut-ligation mix	
component	amount
backbone	1 $\mu$ l
A5	1 $\mu$ l
5B	1 $\mu$ l
BC	1 $\mu$ l
10x T4 DNA Ligase Reaction buffer	2 $\mu$ l
T4 DNA Ligase (Affymetrix)	1 $\mu$ l
BbsI (NEB)	2 $\mu$ l
ddH2O	11 $\mu$ l

cut-ligation program		
temperature	time	cycles
37 °C	5 min	20
16 °C	5 min	
50 °C	5 min	1
80 °C	5 min	1
4 °C	$\infty$	1

## 3.5 Protein analysis techniques

### 3.5.1 *In vitro* transcription and translation

In order to test the catalytic activity of the nucleases an *in vitro* cut assay was developed. The proteins were generated using the TNT Quick Coupled Transcription/Translation System (Promega). Protein coding sequences were under control of a T7 promoter. The following reaction was set up and incubated for 90 minutes at 30 °C.

transcription / translation mix	
component	amount
DNA	2 $\mu$ l (1 $\mu$ g)
TNT Quick Master Mix	20 $\mu$ l
methionine (1mM)	0.5 $\mu$ l
ddH <sub>2</sub> O	2.5 $\mu$ l

### 3.5.2 Preparation of cell extracts

Cell extracts were generated by washing  $5 \times 10^7$  HEK293 cells in PBS and lysing them in 1 ml RIPA buffer for 20 minutes on ice. In case protein lysates were further used for affinity chromatography and EMSA protease inhibitor was added. Then the lysate was sonified using a Branson sonifier 250-D (10 sec on / 50 sec off, 1 minute, 30 % amplitude, on ice). Cell debris was pelleted by centrifugation for 15 minutes at 20000 g, 4°C. The supernatant was transferred to a new tube.

For SDS- analysis the protein concentration was measured with Bradford solution (Bio-Rad) in a BioPhotometer (Eppendorf).

### 3.5.3 Strep-Tactin affinity chromatography

Proteins were expressed with a tandem strep tag and purified by Strep-Tactin affinity chromatography. Therefore, six to eight 15 cm cell culture dishes were transfected with the expression plasmid coding for the desired protein (see 3.2.3). After 48 hours, the cell extracts were prepared in 10 ml RIPA (see 3.5.2). Washing and elution buffers were supplemented with protease inhibitor. All steps were performed at 4 °C. A poly-prep chromatography column was loaded with 300  $\mu$ l of Strep-Tactin sepharose beads and equilibrated with 300  $\mu$ l washing buffer. The protein lysate was loaded onto the column 3 times followed by 5 washing steps with 300  $\mu$ l washing buffer each. The protein was eluted in 6 fractions 150  $\mu$ l each. The protein concentration was first measured with Bradford solution and then verified on a SDS gel in comparison to different amounts of BSA.



### 3.5.4 Electrophoretic mobility shift assay (EMSA)

Protein-DNA interaction was analyzed using electrophoretic mobility shift assays. The assay is based on altered migration behaviour of a protein-DNA complex in comparison to unbound DNA in a polyacrylamide gel (FRIED and CROTHERS 1981). Proteins were purified using Strep-Tactin affinity chromatography (see 3.5.3). As DNA templates PCR products or oligo nucleotides were used.

Two complementary strands of oligo nucleotides were annealed by incubating the following mix for 10 minutes at 95 °C. The Thermoblock was switched off until temperature dropped to room temperature level.

annealing reaction	
component	amount
forward oligo (100 µM)	20 µl
reverse oligo (100 µM)	20 µl
NaCl (5M)	1.64 µl

Annealed oligo nucleotides were loaded onto a 10 % native polyacrylamide gel and after running the gel extracted like described in chapter 3.4.4 in a volume of 100µl. PCR template was generated by setting up two PCR reactions (3.4.6) and purifying the product from a 1 % agarose gel. The PCR product was eluted in 30 µl and concentration was measured using the nano drop instrument.

The DNA fragments were radioactively labelled in a T4 polynucleotide kinase reaction (T4-PNK) with [ $\gamma$ -<sup>32</sup>P] as the phosphate group donor. Therefore, the following reaction was set up and incubated at 37 °C for 1 hour and then for 10 minutes at 65 °C.

T4-PNK reaction	
component	amount
DNA	5 µl (oligo) / 28 µl (PCR)
10x buffer A (Fermentas)	2 µl
T4-PNK (Fermentas)	2 µl
[ $\gamma$ - <sup>32</sup> P] dATP (30 mCi/ml)	3 µl
ddH <sub>2</sub> O	ad 50 µl

The labelled fragments were purified using G50-Sephadex columns (GE Healthcare) and radioactive emission was measured.

Protein and DNA was incubated for 10 minutes at room temperature in the following mix. Serial dilutions were done from 1:2 to 1:256 in dilution buffer before adding DNA.

EMSA reaction	
<b>component</b>	<b>amount</b>
Protein	6 $\mu$ l
10x binding buffer	1 $\mu$ l
poly (dI/dC) (1 $\mu$ g/ $\mu$ l)	0.5 $\mu$ l
DNA template	5000 cpm
ddH <sub>2</sub> O	ad 10 $\mu$ l

After incubation, glycerol was added to the samples at a final concentration of 5 % and the samples were loaded onto a 10 % native acrylamide gel, which had been prerun for 30 minutes. Additionally to the 9 samples (undiluted + 8 dilution steps), in the last lane DNA alone in dilution buffer with 5 % glycerol and bromophenol blue was loaded. The gel was run at 4 °C and 120 V until the bromophenol blue dye reached the end of the gel. The gel was incubated for 15 minutes in fixing solution and then dried in a vacuum-gel-dryer on Whatman paper.

DNA was visualized using a High performance autoradiography film (GE Healthcare). For densitometric quantification of protein-DNA complexes an image plate (Fujifilm) was exposed and read-out with a Phosphoimager (Fujifilm). Quantification of radioactive emission was done using the program AIDA (Raytest).

### 3.5.5 Determination of the equilibrium dissociation constant ( $K_d$ -value)

The  $K_d$ -value defines the protein concentration needed to achieve a half-maximum binding at equilibrium, expressed here in nM concentration (RYDER *et al.* 2008). The  $K_d$ -value was experimentally assessed using electrophoretic mobility shift assays (3.5.4). The amount of protein used in ng was converted to nmol

([http://molbiol.ru/eng/scripts/01\\_04.html](http://molbiol.ru/eng/scripts/01_04.html)) and the ratio of protein-DNA complex to unbound DNA was calculated. The  $K_d$ -value was calculated using the program Prism6 (Graphpad). The bound fraction (linear scale) was plotted against the protein concentration (logarithmic scale) and the  $K_d$ -value was calculated using the equation “One site – Specific binding with Hill slope”.

### 3.5.6 SDS-PAGE

SDS-PAGE was performed using a Mini Protean Tetra Cell (Bio-Rad). Acrylamide gels were cast as described in chapter 2.8. Samples were prepared as described below and heat denatured for 5 minutes at 100°C in a sand bath.

SDS-PAGE sample	
component	amount
protein	40 µg (Western blot) 250 ng -1000ng (Coomassie staining)
Laemmli buffer (4x)	5 µl
ddH <sub>2</sub> O	ad 20 µl

Proteins were separated at 180V until the bromophenol blue dye migrated out of the gel.

### 3.5.7 Western blot

After running a SDS gel the proteins were transferred onto a Hybond ECL membrane (GE Healthcare) using a Mini Trans-Blot Cell (Bio-Rad). The transfer was performed at 100 V for 80 minutes. After disassembly the membrane was incubated for 30 minutes in blocking solution at room temperature and then with the primary antibody, diluted in blocking solution, at 4 °C overnight. The next day, the membrane was washed five times for 5 minutes each in PBS-T and incubated with the HRP-conjugated secondary antibody, diluted in blocking solution, for 45 minutes at room temperature. The membrane was again washed 5 times for 5 minutes and for a few seconds in PBS and afterwards ddH<sub>2</sub>O.

After incubating the membrane with ECL for 5 minutes the bands were detected using a x-ray film. Antibodies, which were used are listed in chapter 2.3.

### 3.6 *In vitro* cut assay

Proteins were *in vitro* translated as described in chapter 3.5.1 and used immediately. All proteins were fused to a Flag tag to allow a comparison of expression levels via Western blot analysis (chapter 3.5.6/7). From each lysate 2  $\mu$ l were used. DNA templates were generated by amplifying the GFP coding sequence via PCR (3.4.5) and purifying the product from a 1 % TBE agarose gel (3.4.3). The following reaction was set up and incubated for 90 minutes at 37 °C.

<i>in vitro</i> cleavage reaction	
<b>component</b>	<b>amount</b>
DNA	200 ng
10x buffer 4 (NEB)	2 $\mu$ l
NaCl (1M)	2 $\mu$ l
BSA (NEB)	0.2 $\mu$ l
nuclease 1	2 $\mu$ l
nuclease 2	2 $\mu$ l
ddH <sub>2</sub> O	ad 20 $\mu$ l

The reaction was separated on a 1.2 % TAE gel. Quantification in 5.6.1 was performed using the Quantum ST4 software (Vilber Lourmat).

### 3.7 *In vivo* cut assay

To assess and compare the *in vivo* activity of different nucleases U-2 OS cells were used, which harbor a genomically integrated copy of *gfp*. Cells (1x10<sup>6</sup> per transfection) were transfected with 3 plasmids (2  $\mu$ g each), using the Neon transfection device from Invitrogen (3.2.4).

- i) p4700, coding for mCherry as a transfection control
- ii) nuclease 1 expression construct
- iii) nuclease 2 expression construct

After transfection, the cells were cultured for 48 hours and analyzed using a LSR Fortessa instrument (Beckton Dickinson) (chapter 3.2.7). The cells were first gated for mCherry to restrict the analysis to transfected cells. The mCherry-positive population was then examined for GFP-positive and GFP-negative cells.

### 3.8 Measurement of the DNA damage response (DDR)

The DNA damage response caused by the expression of different nucleases was assessed in K562 cells. The cells were counted and  $5 \times 10^5$  cells transfected with three plasmids using the Neon transfection device (Invitrogen) (chapter 3.2.4).

- i) p4700, coding for mCherry as a transfection control (1  $\mu$ g)
- ii) nuclease 1 expression construct (6  $\mu$ g)
- iii) nuclease 2 expression construct (6  $\mu$ g)

As a negative control nuclease 1 and 2 were replaced by 12  $\mu$ g of an empty expression vector. Etoposide was used as a positive control by incubating cells with 500  $\mu$ M for two hours. After transfection, the cells were cultivated for 48 hours and then stained with an antibody detecting the phosphorylated histone H2A.X (anti-P-histone H2A.X, Cell Signaling) using the Fix&Perm Cell Permeabilization Kit (Invitrogen). Histone H2A.X is phosphorylated following DNA double-strand breaks (DSBs) since it is required for the accumulation of DNA damage response proteins (YUAN *et al.* 2010). Phosphorylation occurs on Ser139 at sites of DNA damage (ROGAKOU *et al.* 1999).

## **4. Results I – Biochemical characterization of DNA binding attributes of ZFNs and TALEs**

ZFNs and TALENs are widely used to modify the genomes of stem cells and somatic cells. Beside their application in genome engineering, not much is known about possible genotoxic effects of ZFNs and TALENs and their biochemical characteristics. This part of the thesis is dedicated to the analysis of several biochemical aspects of both ZFNs and TALENs. I wanted to determine the dissociation constant  $K_d$  of several natural or synthetic ZFNs and TALEs. TALE proteins and their sequence-specific binding characteristics have been discovered only recently. Published  $K_d$  values for TALEs diverge substantially. In order to determine the  $K_d$  values I employed electrophoretic mobility shift assays (EMSAs). Based on the results of an EMSA one can calculate the  $K_d$  value of a protein, which corresponds to the protein concentration necessary to bind 50 % of the DNA template providing a measure of the binding affinity of a protein to its cognate DNA sequence.  $K_d$  values of two ZFNs, a synthetic TALE and the natural TALE HAX3 were assessed. Additionally, the effect of the DNA template length on the  $K_d$  value was analyzed as a second parameter. I also considered the dimerization characteristics of a ZFN pair and the cooperation of the two single proteins on their cognate DNA template.

### **4.1 Generation of *gfp*-specific TALENs**

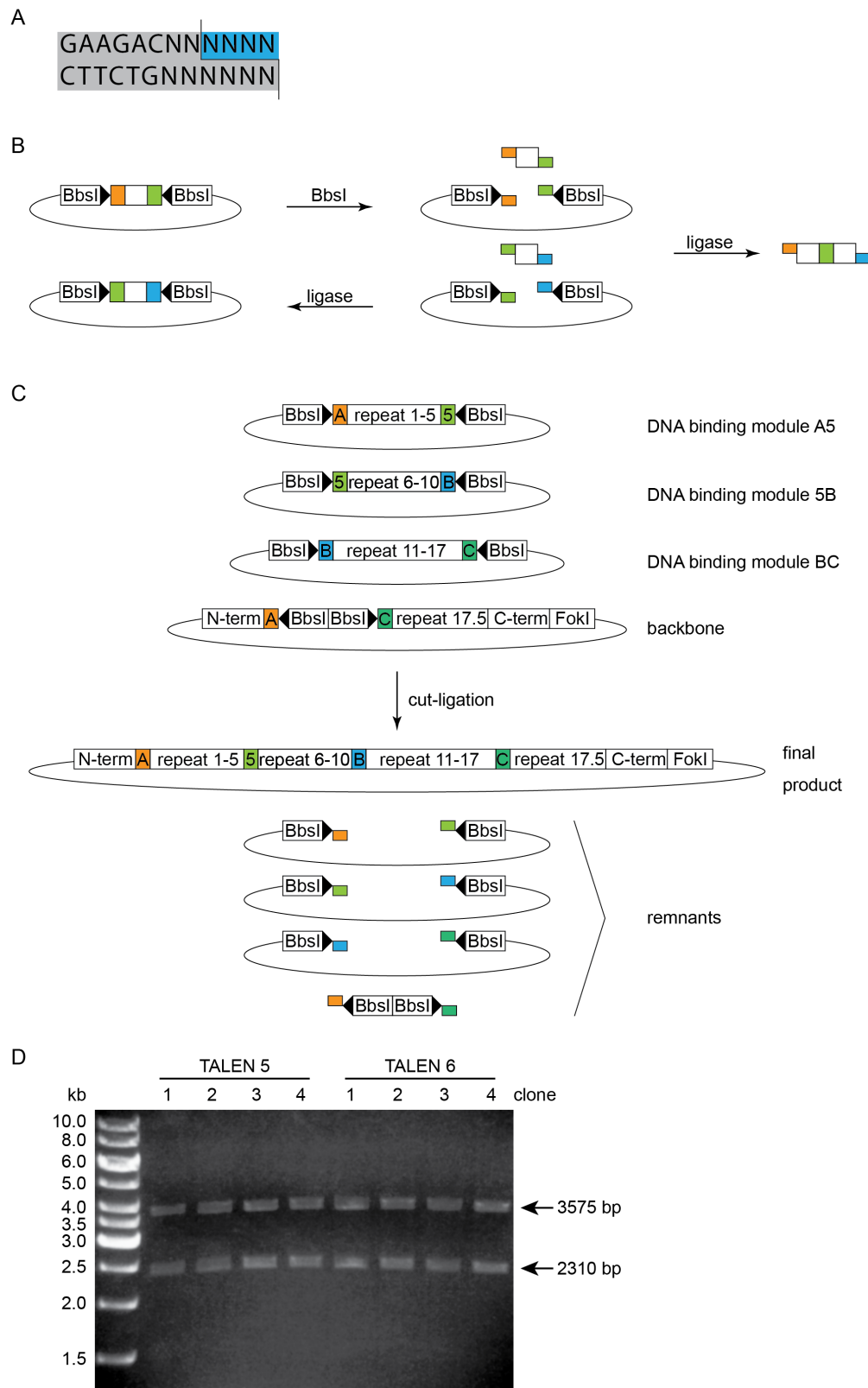
In the framework of this thesis two pairs of TALEN were generated in cooperation with Thomas Lahaye and Orlando de Lange (LMU). A modular cloning approach, exploiting the cleavage characteristic of the type IIS restriction enzyme BbsI, was used to combine three modules encoding the DNA binding domain of the TALEN and one module encoding the TALEN backbone. BbsI cleaves outside its cognate sequence generating a 4

bp single-stranded overhang. This feature allows to generate 256 ( $4^4$ ) unique single-stranded 5' ends, constituting a modular cloning system, which is based on the assembly of multiple fragments in a desired order and orientation (Fig. 4.1 A).

The cloning reaction, called cut-ligation, combines the cleavage by the BbsI restriction enzyme and ligation by the T4 DNA ligase in a single reaction. The BbsI sites bracket the cloning modules and generate specific overhangs upon cleavage (Fig. 4.1 B; small orange, green, and blue boxes indicate three distinct overhangs). Subsequently, two possible ligations can occur in principle. The modules can be religated with their parental plasmids restoring the initial situation. The religation will regenerate the BbsI sites and a second cleavage event will take place. The second possible ligation fuses the two cloning modules in the desired order via their complementary overhangs (Fig. 4.1 B; green box on the right). In this situation the BbsI cognate sequence will not be regenerated excluding a further cleavage and resulting in a stably ligated DNA fragment.

The reaction was set up with the three plasmids encoding the three separate DNA binding domain modules and one plasmid encoding the TALEN backbone (Fig. 4.1 C). The BbsI sites in the plasmids encoding the DNA binding domain modules and the plasmid coding for the TALEN backbone are orientated in opposite directions. Cleavage by BbsI results in the release of fragments encoding the DNA binding domain modules. In contrast, BbsI cleavage of the TALEN backbone encoding plasmid results in an opening of the backbone between the N-terminus and terminal half repeat generating specific overhangs (Fig. 4.1 C). This configuration allows cloning of the three DNA binding modules into the TALEN plasmid backbone. During cut-ligation the DNA binding domain modules and the TALEN backbone are assembled in the correct order. The three plasmid backbones, which harbored the binding domain modules and the insert of the TALEN backbone encoding plasmid remain as left overs (Fig. 4.1 C).

The integrity of the sequences coding for the TALENs after the cut-ligation reaction was verified via a restriction enzyme digestion. The result of assembling TALEN 5 and TALEN 6 is shown here as an example. In both cases all clones showed an intact TALEN encoding gene after performing a control digest (Fig. 4.1 D).



**Fig. 4.1 TALENs 5, 6, 7, and 8 were generated by combining three modules encoding the DNA binding domain and the TALEN backbone module in a cut-ligation reaction**

**A.** The cut-ligation system exploits the cleavage characteristics of the type IIS restriction enzyme BbsI. It cleaves distal to its cognate sequence without nucleotide preference and generates four-base long 5' single-stranded overhangs. Thus, BbsI can produce 256 ( $4^4$ ) different overhangs facilitating the establishment of a modular cloning system with which multiple fragments are assembled in a predetermined desired order and orientation.



**B.** Overview of a cut-ligation with two fragments. On the left, two plasmids are shown each comprising a single cloning module bracketed by BbsI sites. The flanking BbsI sites were designed to generate specific overhangs after BbsI cleavage (middle) illustrated by two different colors. The two plasmids are cleaved and the resulting fragments assembled in a single reaction with BbsI and T4 DNA ligase. Two different ligations can occur in principle. (i) A binding module and its parental plasmid are religated, which results in the initial situation (left). The plasmid will be cut again because the two BbsI sites are restored. (ii) The two free modules are ligated together via their complementary (green) overhangs (right). After ligation the previous BbsI cleavage sites form a sequence that is devoid of the BbsI site. As a consequence, BbsI cannot cleave within the newly formed junction (green), which results in a stably fused fragment (right).

**C.** The cut-ligation reaction for TALEN generation was performed with three plasmids containing the modules A5, 5B and BC and a fourth plasmid with the TALEN backbone. They yield the TALEN DNA binding domain and the TALEN backbone with the TALEN N-terminus and C-terminus and the FokI nuclease domain. The BbsI sites are orientated in opposite directions in the binding modules of donor plasmids (A5, 5B and BC) compared to the TALEN backbone donor plasmid as indicated by the arrow heads. As a consequence, BbsI cleavage releases the binding modules from the plasmid whereas BbsI cleavage in the backbone donor plasmid opens it between the N-terminus and the last half repeat (repeat 17.5) conferring the insertion of the binding modules into the backbone. After cut-ligation, the three DNA binding modules were cloned in the correct order and orientation resulting in an intact TALEN gene in the backbone plasmid.

**D.** After cut-ligation four clones of each TALEN construction were analyzed by a restriction enzyme digestion with XmnI. To illustrate the efficiency of the strategy the ethidium bromide stained agarose gel shows a total of eight TALEN 5 and 6 clones. After restriction digest each clone delivered fragment sizes of 3575 bp and 2310 bp as expected.

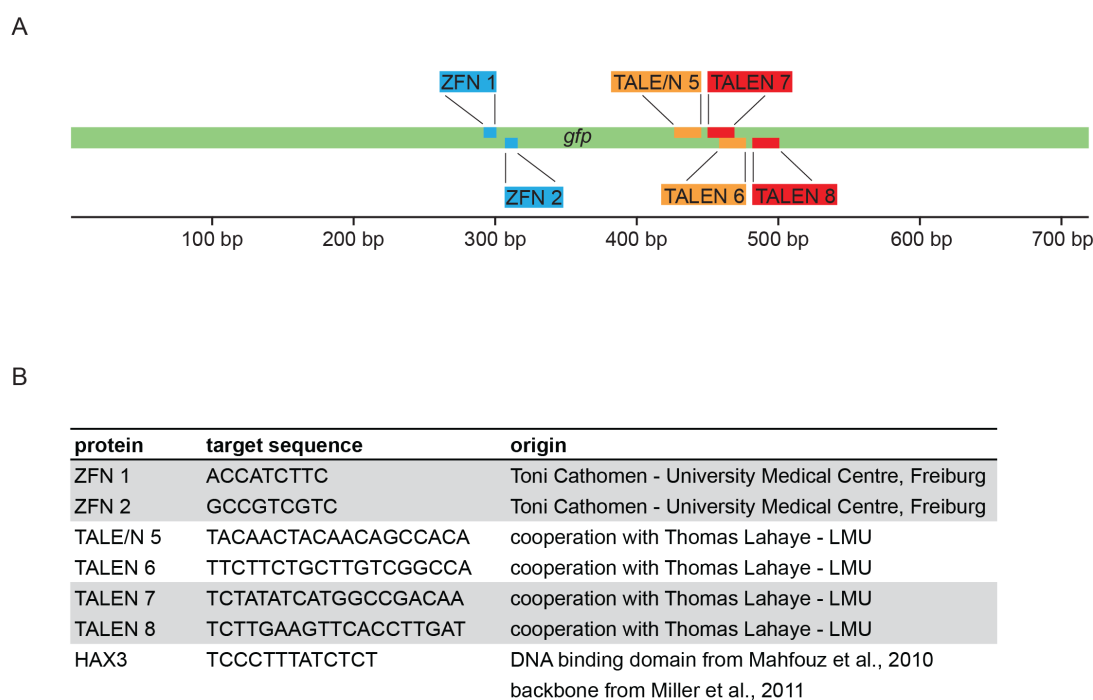
Figure adopted from Morbitzer et al., 2011.

## 4.2 Overview of TALE/Ns and ZFNs used in this thesis

In my thesis different TALE/Ns and ZFNs were used for initial biochemical analyses (Chapter 4) and antiviral effector nucleases (Chapter 5). Here, the proteins and their cognate binding sequences are summarized (Fig. 4.2). ZFN 1 and ZFN 2 were kindly provided by Toni Cathomen (University Medical Centre, Freiburg). The ZFNs target a sequence of 9 bp each separated by a 6 bp spacer in the *gfp* coding sequence. Both ZFNs were used for biochemical studies as well as a positive control in DNA cleavage experiments in Chapter 5. The cut-ligation reaction, described above, was used to generate two TALEN pairs that target sequences within the *gfp* coding region in the cleavage experiments in Chapter 5. The cognate sequence of a single TALEN has a length of 19 bp. One bp (always a T residue) is targeted by the N-terminus of the protein, 17 bp are targeted by the repeats in its DNA binding domain and one bp is targeted by the terminal half repeat, resulting in a total target sequence motif of 19 bp. The TALEN pair one consists of TALEN 5 and 6. A twelve bp spacer separates the target sequences of both TALENs. The TALEN pair two consists of TALEN 7 and 8. A 13 bp spacer separates their target sequences (Fig. 4.2). In an additional cloning reaction TALE 5 was

constructed. It shares the DNA target sequence with TALEN 5 but lacks the FokI cleavage domain and was used in biochemical analyses, only.

The target sequences of the proteins are summarized in Figure 4.2. In addition, the natural TALE HAX3 was used for biochemical studies in this thesis. HAX3 and its cognate sequence, which is not part of the *gfp* coding sequence is not depicted in Figure 4.2 A, but listed in Figure 4.2 B.



**Fig. 4.2 Overview of the localization and target sequences of the different TALE/Ns and ZFNs used in this study**

**A.** Two ZFNs (blue bars) and four of the TALE/Ns (orange and red bars) target different but adjacent DNA sequences within the coding region of *gfp* (green bar). ZFN 1 and 2 (blue bars) target *gfp* about 300 bp downstream of the start of the open reading frame. ZFN 1 targets the top strand whereas ZFN 2 targets the bottom strand. TALEN 5 and 6 (orange bars) and TALEN 7 and 8 (red bars) target the *gfp* sequence between 400 bp and 500 bp downstream of the start of the open reading frame. TALEN 5 and 7 target the top strand whereas TALEN 6 and 8 target the bottom strand. In addition, TALE 5 was constructed to facilitate the assessment of the biochemical binding characteristics of artificially constructed TALEs. TALE 5 is identical to TALEN 5 but lacks the FokI nuclease domain.

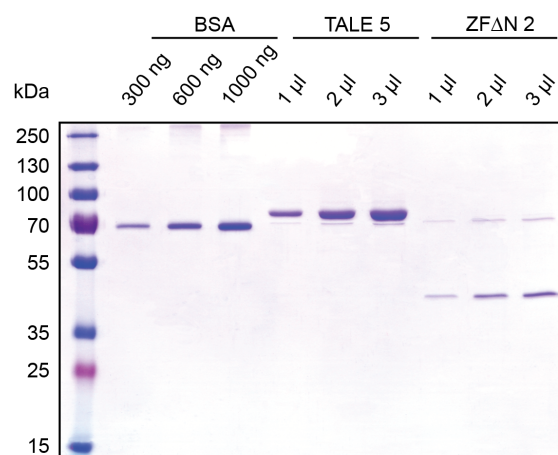
**B.** The table summarizes the proteins used in this thesis and depicts their target sequences and origins. In addition to the proteins shown in panel A the TALE HAX3 is included, which is a natural TALE identified in *Xanthomonas*. It was used in biochemical binding studies in comparison with the artificially engineered TALE 5.

TALE= transcription activator-like effector; TALEN= transcription activator-like effector nuclease; ZFN= zinc finger nuclease

### 4.3 Protein purification with HEK293 cell lysates

For the usage in electrophoretic mobility shift assays (EMSAs) the proteins had to be synthesized and purified. HEK293 cells were transiently transfected with expression plasmids encoding the proteins of interest fused to a tandem Strep-tag. 48 hours after transfection cells were lysed and proteins purified by Strep-Tactin affinity chromatography. The protein concentration was assessed by Bradford measurements (not shown). To assess the purity of the protein lysate and verify the concentration measured by Bradford different amounts of purified protein and BSA were loaded on a SDS-PAGE gel, which was stained with Coomassie brilliant blue.

Figure 4.3 shows an example depicting the purified proteins TALE 5 and ZFΔN 2 (ZFN 2 with inactivated FokI nuclease domain). Both proteins show a high purity and a comparison to different amounts of BSA verify the Bradford measurement results (Fig. 4.3).



**Fig. 4.3 Purification of HEK293 cell lysates containing ectopically expressed TALE 5 or ZFΔN 2 proteins by affinity chromatography**

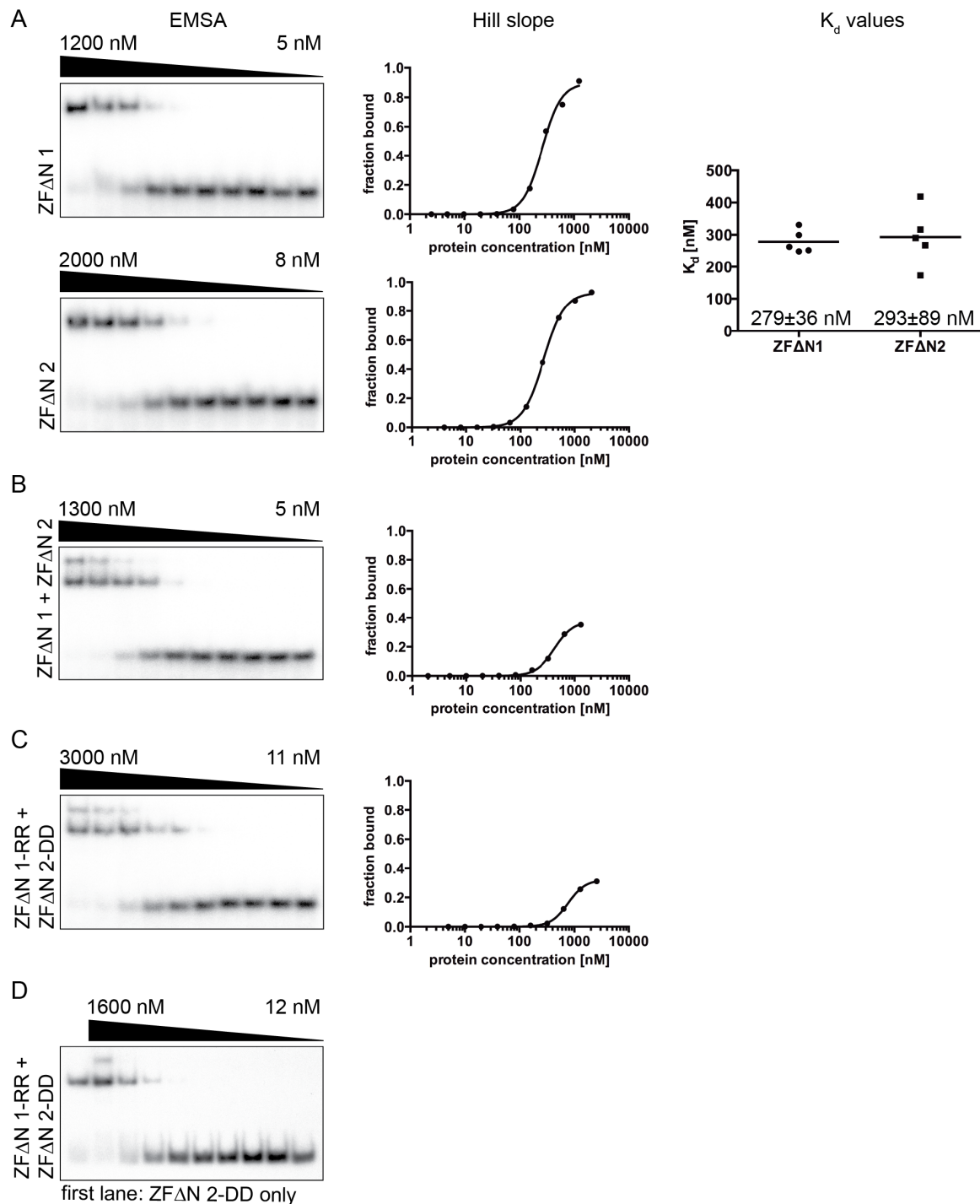
HEK293 cells were transiently transfected with expression plasmids encoding TALE 5 or ZFΔN 2. The cells were lysed and the proteins were purified prior to the EMSA experiments by affinity chromatography. The purity and concentration of the proteins were analyzed on a SDS-gel stained with Coomassie brilliant blue shown here. Different amounts of BSA were loaded to assess the concentration of the purified proteins. The results with TALE 5 and ZFΔN 2 are shown here as two examples. Both proteins achieved a high purity. Bradford measurements (not shown) were in a comparable range (TALE 5= 300 ng/μl; ZFΔN 2= 200 ng/μl) confirming the signals visualized in the Coomassie-SDS-Page gel.

BSA = bovine serum albumine

## 4.4 Electrophoretic mobility shift assays with ZFΔNs

A dimerization of ZFNs via their FokI nuclease domain results in the cleavage of DNA and a double-strand break. DNA cleavage would be detrimental in EMSA experiments, because the radioactively labelled DNA template would be fragmented preventing protein binding and the calculation of the  $K_d$  values of the proteins. For this reason the catalytic activity of the FokI nuclease domain was inactivated by a single alanine substitution of Asp-450 in the the catalytic center (WAUGH and SAUER 1993). In the following, ZFNs with an inactivated FokI nuclease domain are denoted ZFΔN. First, the  $K_d$  values of single ZFΔN monomers were assessed. Both ZFΔN 1 and ZFΔN 2 showed comparable  $K_d$  values of  $279 \pm 36$  nM and  $293 \pm 89$  nM, respectively (Fig 4.4 A). Next, both ZFΔNs were combined in EMSA experiments to assess the binding characteristics of a dimer consisting of ZFΔN 1 and ZFΔN 2, which resulted in the appearance of two additional bands (Fig. 4.4 B, C, D). The faster migrating shift bands resemble a DNA fragment bound by a single ZFΔN molecule. The top most band most likely is the trimeric complex consisting of ZFΔN 1, ZFΔN 2 and the DNA molecule. Comparing the intensities of the upper and middle bands suggests a low stability of the trimeric complex as a minority of DNA fragments is bound by a ZFΔN dimer, only. The Hill slope analysis reveals that about 40 % of the DNA fragments are bound by a dimer at the highest protein concentration. As a consequence, the calculation of the  $K_d$  value is not possible as a nearly complete complex formation is necessary (Fig. 4.4 B).

Single amino acid substitutions in the FokI domain can increase the cleavage specificity of a ZFN pair by preventing homodimerization of ZFN monomers and favoring the formation of heterodimers. Here the DNA binding characteristics of ZFΔN 1-RR and ZFΔN 2-DD was assessed (Fig. 4.4). The EMSA shows that the upper band, reflecting the trimeric complex, is less prominent compared to ZFΔN 1 and 2 at the same concentration of proteins indicating an even weaker complex formation of the mutated FokI domains (compare Fig. 4.4 C and D). Again, a calculation of the  $K_d$  value was not possible because optimal saturation levels were not reached (Fig 4.4 C).



**Fig. 4.4 ZFΔN 1 and 2 preferentially bind their target sequence as monomers with a  $K_d$  value of approximately 300 nM**

Biochemical binding characteristics of ZFN proteins to their target sequences were assessed by EMSA experiments. ZFNs with inactivated FokI nuclease domains (ZFΔN) were used. They were expressed in HEK293 cells, purified by affinity chromatography and incubated with oligo nucleotides encompassing the cognate target sequences of the zinc finger proteins. On the left the results of representative EMSAs are shown. The faster migrating bands indicate free, radioactively labelled DNA fragments. DNA fragments bound by ZFΔN proteins migrate slower and near the top of the EMSA gel. In the middle panels the bound fraction of DNA fragments is plotted as a function of the protein concentration. The Hill slope equation was used to calculate the  $K_d$  values. In the right panel  $K_d$  values are plotted to summarize the results.

**A.** EMSAs with ZFΔN 1 or 2 show that both proteins bind to their target sequences with increasing protein concentrations (left panels). The Hill slopes reveal that in both cases saturation was reached and 50 % of

the DNA fragments were bound at an approximate protein concentration of 300 nM (middle panel). The right panel shows the  $K_d$  values of ZFΔN 1 and 2 from five independent experiments (means and standard deviation).

**B.** When using both ZFΔN 1 and 2 proteins together in EMSAs (upper left panel) the majority of DNA fragments is bound by one ZFΔN protein, only (middle bands). A small fraction of the DNA fragments is bound by a ZFΔN dimer (upper bands) indicating a very low stability of the trimeric complex. As a consequence, dimerization of ZFΔNs on DNA appears to be a rare event and instable. At the highest concentration of 1300 nM protein about 40 % of the DNA fragment are bound by both, ZFΔN 1 and 2 (upper right panel). This value is too low to calculate a  $K_d$  value for the binding of ZFΔN 1 and 2 dimer to its target sequence but suggests that the  $K_d$  value will exceed 1.3  $\mu$ M.

**C.** The EMSA was also performed using ZFΔNs with mutations in their FokI sequences, which allow heterodimer formation, only, in order to increase their target specificity. The EMSA shows an even weaker dimerization of the mutated ZFΔNs compared to their parental ZFΔNs (left panel). Plotting the bound fraction against the protein concentration indicated that it needs a two-fold higher protein concentration to reach the same fraction of bound DNA fragments compared to the parental ZFΔNs. Again, a calculation of a  $K_d$  value is not possible as the bound fraction is too low.

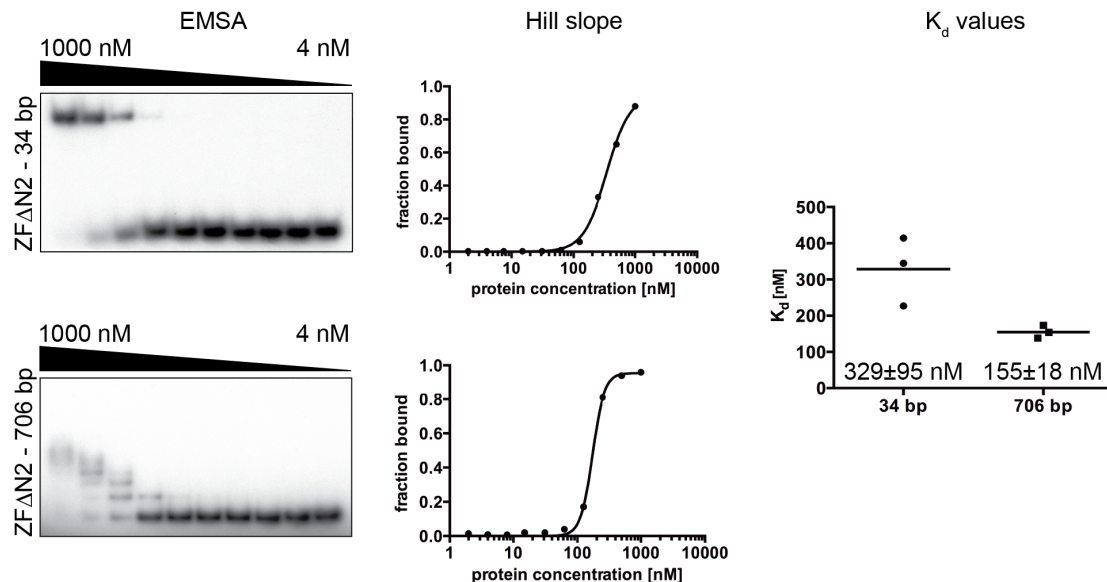
**D.** To verify that the middle bands in panels B and C reflect a DNA fragment bound by a single ZFΔN an EMSA was performed, in which both ZFΔN 1-RR and 2-DD were used, except in lane 1 where ZFΔN 2-DD was used. The EMSA shows one band in the first lane (one ZFΔN bound to DNA), and a second band, which migrates slower in the second lane indicative of the trimeric complex of one DNA fragment bound by both ZFΔN 1-RR and 2-DD. The EMSA clearly shows that the middle band is a complex of a DNA fragment bound by a single ZFΔN.

ZFN=zinc finger nuclease; EMSA= electrophoretic mobility shift assay

Very likely, the faster migrating shift band in Figure 4.4 B and C represents a DNA fragment bound by a single ZFΔN monomer. To analyze the complex situation an EMSA was performed with ZFΔN 2-DD protein in the first lane, only, resulting in a single band. In the adjacent lanes, both ZFΔN 1 and 2 resulted in two shift bands. The EMSA clearly shows that the faster migrating shift band is a complex of DNA fragments and a single ZFN monomer as it migrates similar to the single band in lane one, which is the monomeric protein complex with DNA (Fig 4.4 D).

The length of the DNA fragment used in the EMSAs could influence the binding affinity of ZFNs or other DNA binding proteins (see below). To address this issue, EMSA experiments with ZFΔN 2 were performed with two DNA templates, which differed in length. Template one was a DNA fragment with the length of 34 bp and already used in the EMSA experiments shown in Figure 4.4. Additionally, a DNA template with a length of 706 bp was employed. It was generated by PCR amplifying the target sequence of ZFΔN2 within the GFP encoding ORF. EMSA experiments with the 34 bp DNA template led to a  $K_d$  value of  $329 \pm 95$  nM comparable to the  $K_d$  value of ZFΔN 2 ( $293 \pm 89$  nM ) as assessed in Figure 4.4. EMSA experiments with the 706 bp DNA template resulted in the appearance of additional bands with increasing protein concentration suggesting that multiple ZFΔN2 proteins can bind to a single but extended DNA template. Consequently, the longer 706 bp DNA template yielded a slightly lower  $K_d$  value of  $155 \pm 18$  nM

indicative of a minor impact of the template size on the  $K_d$  value (Fig. 4.5). It is obvious from these experiments, however, that ZF $\Delta$ N2 does show unspecific binding to those DNA sequences that do not encompass its cognate DNA sequence motif.



**Fig. 4.5 ZF $\Delta$ N 2 binds its minimal target sequence efficiently**

The biochemical binding characteristics of purified ZF $\Delta$ N 2 to two DNA target fragments differing in size was assessed in EMSA experiments.

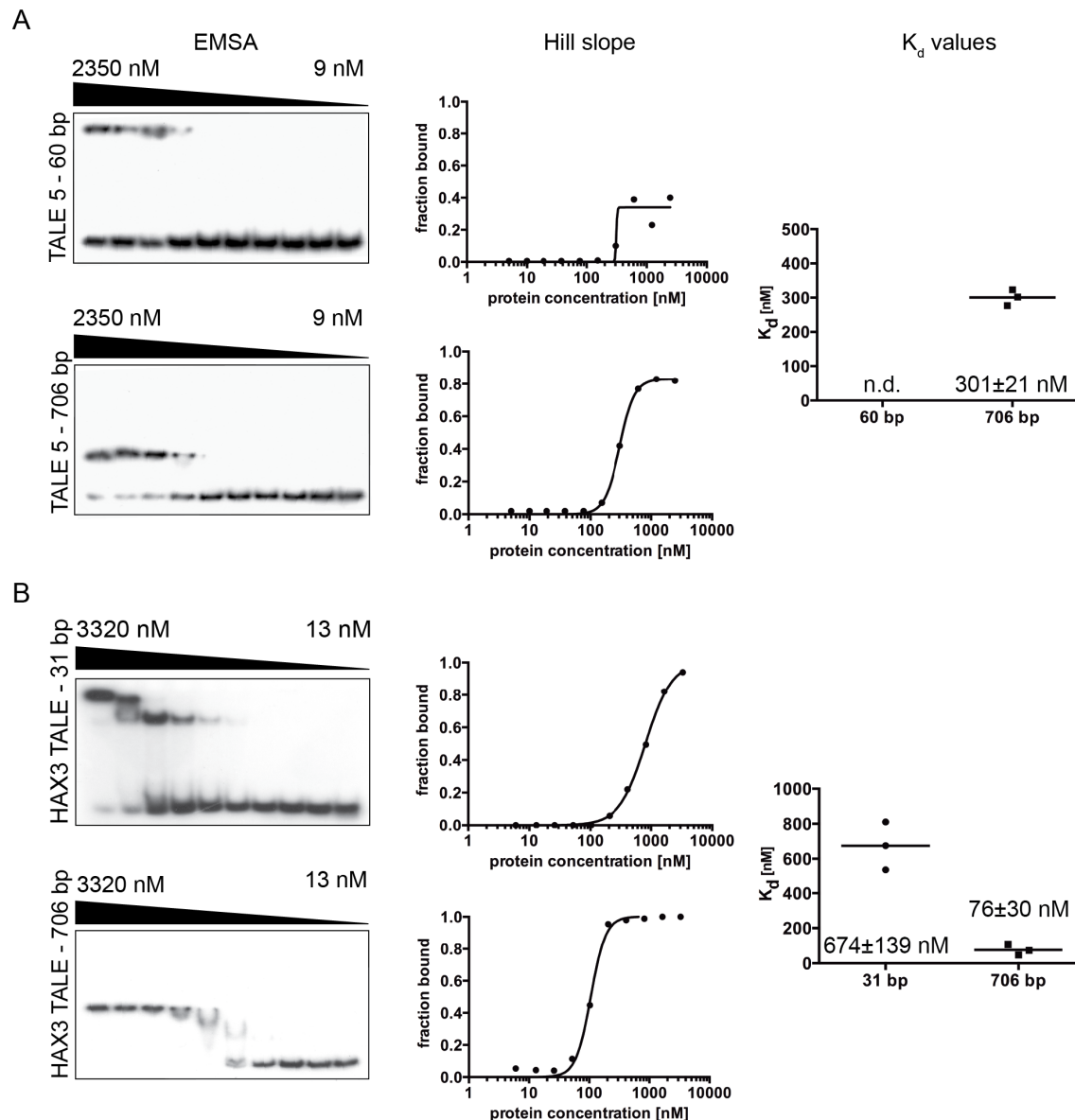
In the upper left panel a template with a minimal length of 34 bp harboring the ZF $\Delta$ N 2 target sequence was analyzed. As shown in Figure 4.4, ZF $\Delta$ N 2 binds to its target sequence as a function of protein concentration. The Hill slope curve shows that half of the DNA templates are bound at a protein concentration of approximately 300 nM (upper middle panel). With a template of 706 bp length, which encompasses one copy of the ZF $\Delta$ N 2 binding sequence additional bands appeared with increasing protein concentrations indicating that several ZF $\Delta$ N 2 molecules can bind non-specifically to this DNA fragment (lower left panel). The Hill slope curve shows a half-maximum binding of ZF $\Delta$ N 2 to DNA at a protein concentration of 150 nM (lower middle panel). The right panel summarizes the results of several independent EMSA experiments comparing ZF $\Delta$ N 2 binding characteristics with the two templates. ZF $\Delta$ N 2 binds to a template with a length of 706 bp with a  $K_d$  value of  $155 \pm 18$  nM as compared to  $329 \pm 95$  nM when a shorter template of 34 bp in length was used.

ZF $\Delta$ N = zinc finger nuclease with inactivated nuclease domain; EMSA = electrophoretic mobility shift assay

## 4.5 Electrophoretic mobility shift assays with TALEs

Biochemical binding characteristics of the *gfp*-specific, synthetically generated TALE 5 and the natural TALE HAX3 proteins were assessed in EMSA experiments. For both proteins two templates differing in length were compared. EMSAs with TALE 5 were performed with a short DNA template with a length of 60 bp and a longer DNA template with a length of 706 bp. The experiments showed that the binding affinity of TALE 5 to the 60 bp DNA template was too low to achieve a saturation of binding required for the

calculation of a  $K_d$  value. Usage of the longer 706 bp DNA template led to a much stronger binding of DNA sufficient to calculate the  $K_d$  value of  $301 \pm 21$  nM (Fig 4.6 A). EMSA experiments with the HAX3 TALE were performed with a short DNA template with a length of 31 bp and a longer DNA template with a length of 706 bp. In both cases a



**Fig. 4.6 Binding affinity of TALE proteins to DNA increases with template length**

Binding characteristics of the synthetic TALE 5 and the natural TALE HAX3 were assessed in EMSA experiments with DNA templates, which encompass a single copy of the target sequence but differ in length.

**A.** EMSAs with the TALE 5 protein were performed with 60 bp (upper left panel) and 706 bp (lower left panel) templates. The two EMSAs demonstrate a higher binding affinity with the longer DNA template. It is almost completely bound at a protein concentration of about 600 nM (lower left panel, lane 3) but the 60 bp template is not bound efficiently even at the maximum concentration of TALE 5 (upper left panel). The Hill slope curve of this experiment with the shorter DNA template indicates insufficient saturation making it impossible to calculate the corresponding  $K_d$  value (upper middle panel). The Hill slope curve of the binding characteristics of the longer template shows a half-maximum binding of about 300 nM of TALE 5. The right panel summarizes the results of six EMSAs. A calculation of the  $K_d$  value of TALE 5 in combination with the shorter 60 bp template was not possible. The combination of TALE 5 with the longer



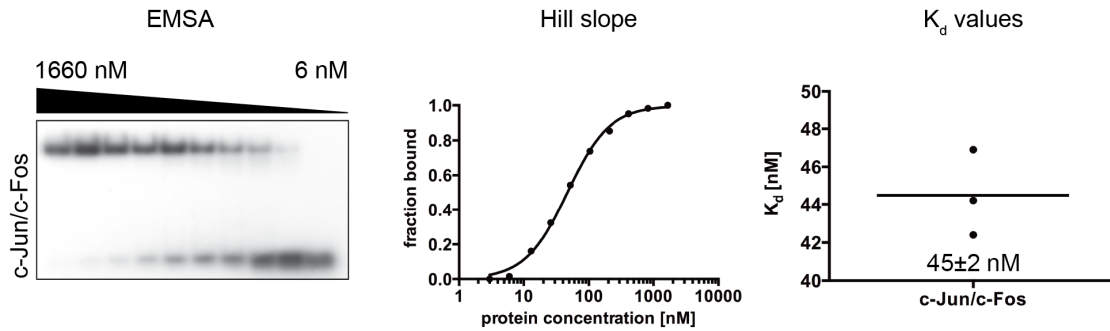
706 bp template resulted in a calculated  $K_d$  value of  $301 \pm 21$  nM suggesting that TALE binding to a minimal DNA binding motive is very inefficient.

**B.** Binding of the natural HAX3 TALE was assessed with a short 31 bp and a long 706 bp DNA template each encompassing a single copy of the HAX3 cognate binding motif. Performing an EMSA with its shorter DNA template shows that saturation of DNA binding was reached at the highest protein concentration of 3320 nM (upper left panel, lane 1) whereas the longer DNA template was completely bound by HAX3 at a protein concentration of approximately 200 nM (lower left panel, lane 5). Again, this TALE prefers binding to a longer DNA template. The corresponding Hill slopes show a half-maximum binding of about 700 nM and 150 nM of protein concentration for the 31 bp and the 706 bp template, respectively. The right panel summarizes the findings indicating that template length plays a crucial role in DNA binding by TALEs. The combination of the HAX3 TALE with the short 31 bp template results in a  $K_d$  value of  $674 \pm 139$  nM whereas the combination of HAX3 with the 706 bp template led to a lower  $K_d$  value of  $76 \pm 30$  nM indicating stronger protein binding to DNA.

$K_d$  value could be calculated. EMSA experiments with the 31 bp DNA template resulted in a high  $K_d$  value of  $674 \pm 139$  nM; the longer DNA fragment resulted in a  $K_d$  value of  $76 \pm 30$  nM indicating a stronger binding affinity of the HAX3 TALE compared to the synthetic TALE 5 (Fig. 4.6 B). Both TALEs bound the longer templates with higher binding affinities in contrast to ZFΔN 2 (Fig. 4.5). This finding suggests that DNA length contributes considerably to the binding affinity of TALE but not ZF proteins indicating different binding mechanics and DNA-protein structures.

## 4.6 Electrophoretic mobility shift assays with the c-Jun/c-Fos dimer

The prototypical AP-1 heterodimer c-Jun/c-Fos was used to validate the EMSA results, which I obtained with ZFΔN and TALE proteins. Towards this end, EMSAs were performed with the purified c-Jun/c-Fos heterodimer and a short synthetic double-stranded DNA template, which harbors the cognate AP-1 sequence motif (5'-TGAGTCA-3'). The resulting  $K_d$  value of  $45 \pm 2$  nM (Fig. 4.7) is comparable to published values of 50 nM (JOHN *et al.* 1996) and  $10.8 \pm 0.7$  nM (CHYTIL *et al.* 1998) confirming the validity of the EMSA results in this study (Fig. 4.4 to 4.6).



**Fig. 4.7 EMSAs with the c-Jun/c-Fos dimer leads to  $K_d$  values in line with the literature**

The  $K_d$  value of c-Jun/c-Fos is known (10-50 nM) and published in 1998 and 1996 by Chytil et al. and John et al., respectively. EMSAs with a purified c-Jun/c-Fos dimer and its cognate binding motif were performed to validate the results obtained in EMSAs with ZFΔN and TALE proteins. The left panel shows one EMSA experiment as an example. Purified c-Jun/c-Fos results in a clear shift (left panel). The Hill slope curve calculation shows a half-maximum binding of c-Jun/c-Fos at a protein concentration of approximately 50 nM (middle panel). The right panel summarizes the results of three experiments revealing a mean  $K_d$  value of  $45 \pm 2$  nM.

EMSA= electrophoretic mobility shift assay

## 5. Results II – TALENs as antiviral factors

The main application of customized nucleases like ZFNs and TALENs is in the field of genomic engineering of stem cells, somatic cells and transgenic organisms like mouse, rat and zebra fish. The aim of genomic engineering is a gene knockout, gene deletion, gene addition, or gene repair. In the framework of this thesis a new application of TALENs was tested. TALENs were generated, that bind within the *gfp* sequence and induce a DNA double-strand break (DSB) upon dimerization. The *gfp* sequence serves as a target sequence in this system to explore synthetic nucleases for their suitability as antiviral factors. In contrast to more common applications in genomic engineering the target sequence in this project was absent in cellular DNA but an integral part of the viral genome to be tested. I asked if TALENs can be engineered that mediate DNA double-strand breaks within the viral sequence inhibiting or blocking the replication of a DNA virus hindering viral *de novo* synthesis.

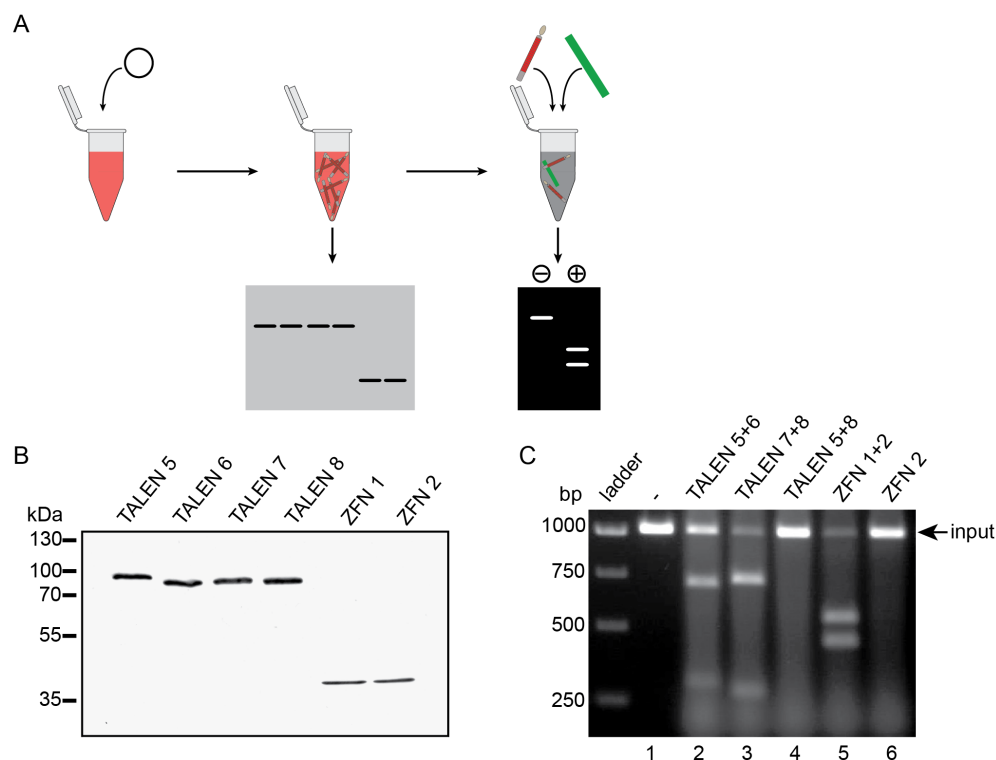
### 5.1 Analysis of the cleavage activity

Two pairs of TALENs were generated as described in the previous chapter 4. The first focus of this project was to assess and to compare the cleavage activity of the two nuclease pairs TALEN 5 and 6 and TALEN 7 and 8. For this purpose, two different cleavage assays were established and performed.

#### 5.1.1 *In vitro* cleavage assay

The first test of cleavage activities of the generated TALENs was performed in a straightforward *in vivo* cleavage assay. All proteins used in the assay were cloned in an

expression plasmid under control of a T7 promoter, which confers protein expression using the TNT Quick Coupled Transcription/Translation System supplied by Promega. The expression levels of the different proteins were assessed by Western blot analysis. The actual cleavage assay was carried out in a restriction enzyme-like reaction. Two *in vitro* expressed nucleases and a PCR product of the *gfp* sequence, harboring the target sequences of all proteins were mixed and incubated. The cleavage results were analyzed on an ethidium bromide stained agarose gel (Fig. 5.1 A).



**Fig. 5.1 Among different TALEN constructs the combination of TALEN 7 and TALEN 8 shows the highest cleavage activity *in vitro***

**A.** Experimental overview: Several expression plasmids that encompass TALEN genes under the control of a T7 promoter (circle) were added to a TNT Transcription/Translation master mix (left panel). The expression levels of the translated proteins (red bars in the vial) were compared by Western blot immunostaining (middle panel). The cleavage activities were assessed in enzymatic reactions with PCR products of the *gfp* sequence as templates (green bar) and different combinations of nucleases (red bar). Reaction products were analyzed on an ethidium bromide stained agarose gel. One band represents the uncleaved *gfp* input DNA (-); two bands indicate a successful cleavage (+).

**B.** Western blot immunostaining indicates comparable expression levels of TALEN 5, 6, 7 and 8 (107 kDa) and ZFN 1 and 2 (41 kDa).

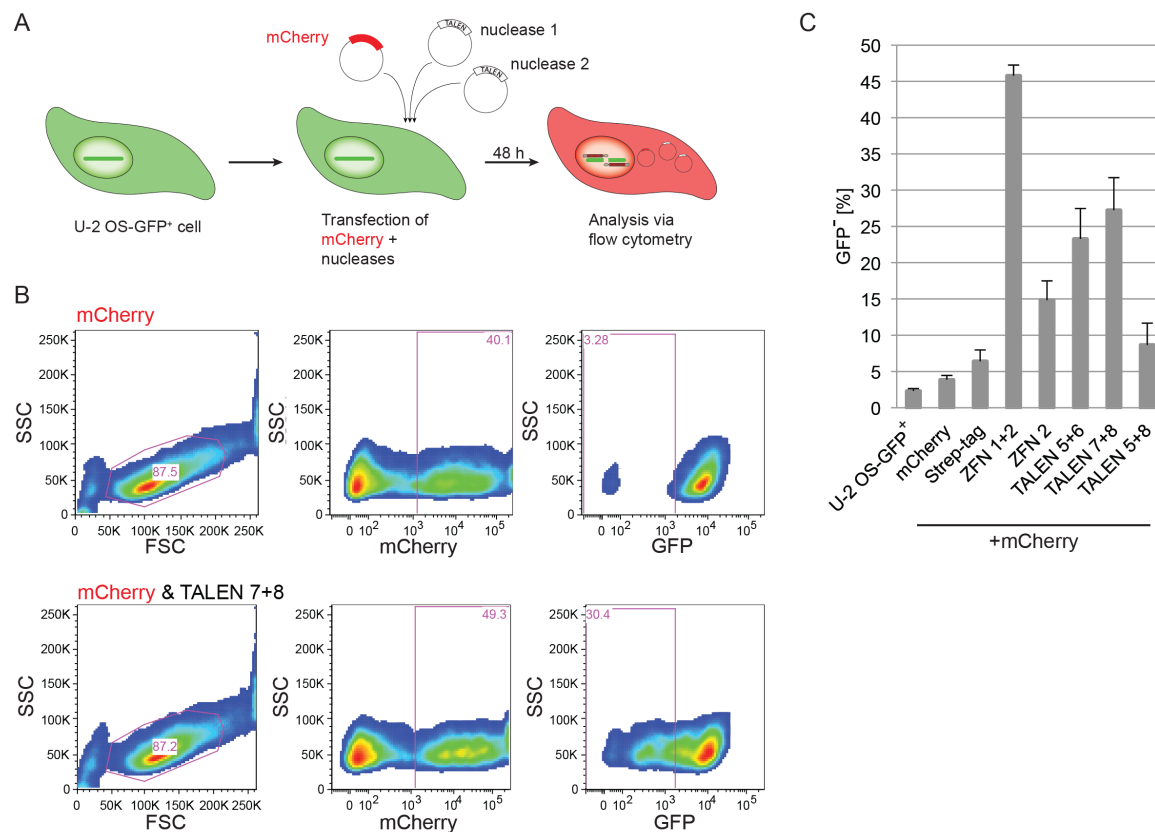
**C.** The ethidium bromide stained agarose gel shows the result of the *in vitro* cleavage assay. Lane 1 shows the *gfp* input DNA as a mock control. Addition of both TALEN 5 and TALEN 6 leads to the appearance of two additional cleavage products but uncleaved input DNA remains (lane 2). Addition of TALEN 7 and 8 shows a higher cleavage activity (lane 3) whereas the addition of both TALEN 5 and 8 shows no cleavage at all (lane 4). The distance between the two recognition sites of TALEN 5 and 8 is too wide, preventing FokI dimerization. The ZFN 1 and 2 pair serves as a positive control and shows a cleavage comparable to TALEN 7 and 8 (lane 5). Addition of ZFN 2 protein alone does not show any cleavage as expected (lane 6). The white background originates from the master mix of the kit used to express the proteins.

ZFN=zinc finger nuclease, TALEN=Transcription activator-like effector nuclease

The Western blot analysis revealed comparable expression levels of the different proteins (Fig. 5.1 B). The agarose gel shows the result of the *in vitro* cleavage assay. Both TALEN pairs had considerable cleavage activities but TALEN 7 and TALEN 8 performed better. The pair consisting of TALEN 5 and TALEN 8 served as a negative control since their target sequences in the *gfp* sequence are too far apart preventing the dimerization of the FokI nuclease domain. Consistent with this notion, this combination did not yield any cleavage products (Fig. 5.1 C, lane 4). ZFN 1 and ZFN 2 served as a positive control since their cleavage activity was already validated by the Cathomen group. The ZFNs showed an activity comparable to TALEN 7 and 8, whereas ZFN 2 alone showed no cleavage activity at all (Fig. 5.1 C, lane 5 and 6).

### 5.1.2 *In vivo* cleavage assay

The *in vitro* cleavage assay showed that the generated TALEN pairs are able to induce a DNA double-strand break. The next step was to evaluate the cleavage activity of the engineered nucleases in cells. A derivative of the U-2 OS cells with a single integrated copy of the *gfp* gene, referred to as U-2 OS-GFP<sup>+</sup>, served as a basis for this experiment. The cells were co-transfected with three expression plasmids encoding mCherry as a transfection control and two nucleases. After 48 hours U-2 OS-GFP<sup>+</sup> cells were analyzed by flow cytometry (Fig 5.2 A). The cells were first gated for the living population and mCherry-positive cells, restricting the analysis to transfected cells, only. Within the mCherry-positive population the GFP-negative and -positive fractions were assessed. After transfection with the mCherry expression plasmid, only, about 3 % of the cells were GFP-negative. Upon co-transfection with expression plasmids encoding mCherry, TALEN 7 and TALEN 8 around 30 % of the cells turned GFP-negative. This result clearly indicates a successful TALEN 7 and 8-mediated DSB within the *gfp* coding sequence (Fig. 5.2 B). Transfection of U-2 OS-GFP<sup>+</sup> cells with mCherry, only, or an additional “empty” expression vector showed no clear difference compared to untransfected cells. The positive control, ZFN 1 and ZFN 2, achieved the highest knockout rate. Upon transfection with ZFN 2, only, a fraction of 15 % of GFP-negative cells was observed. This might be due to unspecific cleavage of the single ZFN. Again, both TALEN pairs showed a considerable cleavage activity in this assay but TALEN 7 and 8 revealed a slightly better



**Fig. 5.2 The combination of TALEN 7 and TALEN 8 shows the highest GFP knockout potential in U-2 OS cells**

Cleavage activities of different TALEN constructs were assessed after their transient expression in U-2 OS cells, which carry a single, chromosomal integrated copy of the targeted *gfp* gene sequence.

**A.** Experimental overview: GFP-positive U-2 OS cells (left panel) were co-transfected with three expression plasmids coding for (i) mCherry as a transfection control (ii) nuclease 1 and (iii) nuclease 2 (middle panel). After transfection, the cells were analyzed by flow cytometry after 48 hours. Cells were gated for mCherry expressing cells to restrict the subsequent analysis to transfected cells, only. In the population of mCherry-positive cells the proportion of cells that had lost GFP expression was assessed (right panel).

**B.** Two examples are shown here to illustrate the analytical workflow. The upper row shows the analysis of cells transfected with the mCherry construct, only, the lower row shows cells transfected with mCherry in combination with TALEN 7 and TALEN 8 constructs. In the first column (left panels) the gates delineate the living cell populations, which were further analyzed. In the second column (middle panels), cells were gated for the mCherry-positive population. In the third column (right panels), the cells were analyzed for GFP expression assessing the proportion of GFP-positive and GFP-negative cells indicating the cleavage of the *gfp* gene. Cells co-transfected with constructs coding for TALEN 7 and TALEN 8 consisted of 30 % GFP-negative cells in contrast to cells transfected with mCherry, only.

**C.** The diagram summarizes the results of cells transfected with different combinations of expression plasmids. Cells transfected with mCherry or an empty expression vector (Strep-tag) show only a minor fraction of GFP-negative cells similar to the untransfected control (U-2 OS). The combination of ZFN 1 and ZFN 2 shows the highest knockout rate with 45 % of GFP-negative cells. ZFN 2 alone shows an obvious loss of GFP expression, indicating unspecific cleavage. The combination of TALEN 7 and TALEN 8 shows an activity exceeding that of TALEN 5 and TALEN 6. The combination of TALEN 5 and 8 has an activity in the range of the negative control transfected with an empty expression construct (Strep-tag) as expected. FSC=forward scatter, SSC=sideward scatter

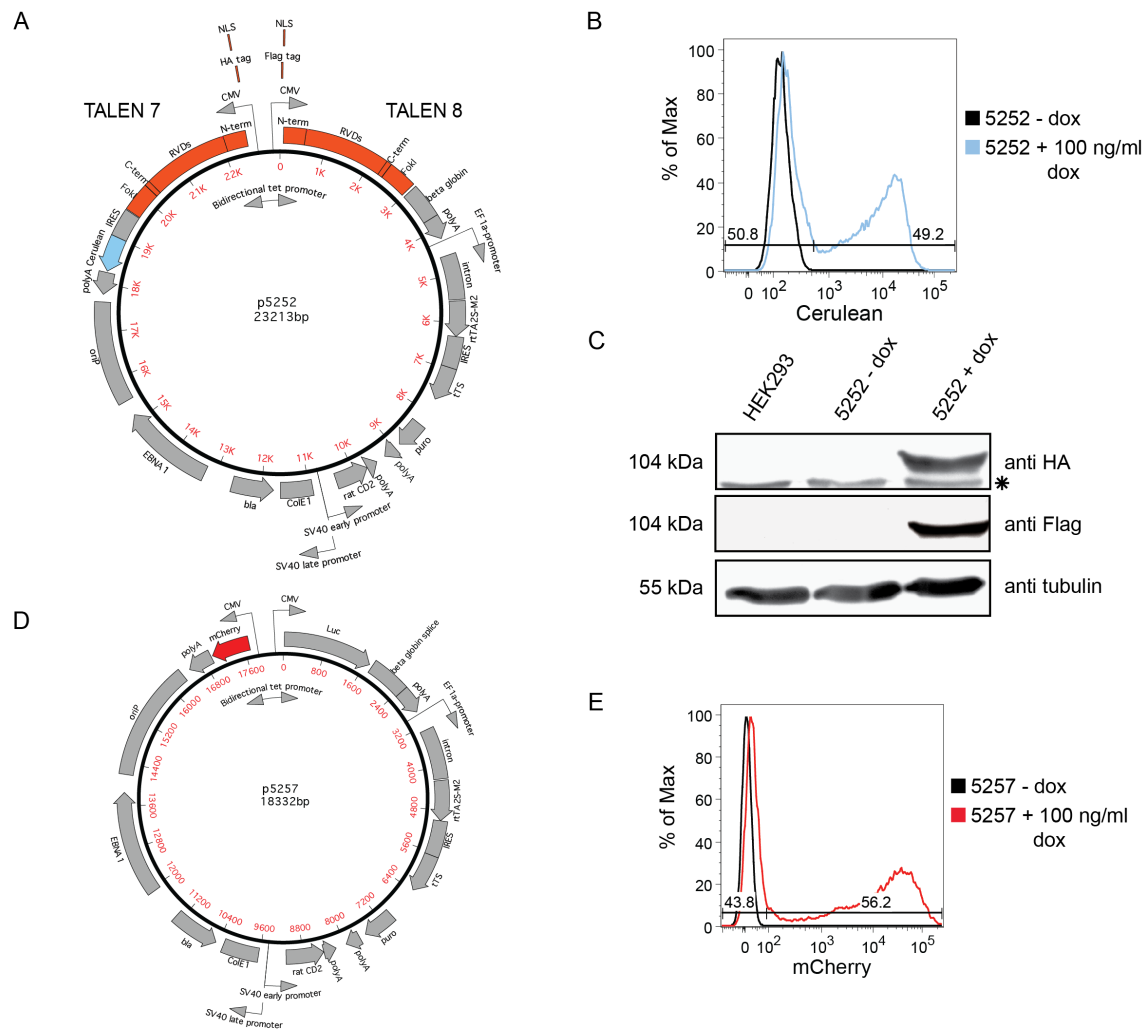
knockout efficiency. The transfection of the cells with TALEN 5 and 8, which should not show any activity, led to GFP-negative cells comparable to the negative control transfected with an empty expression construct (Fig. 5.2 C).

## 5.2 Establishment and characterization of an inducible expression system

The first two cleavage experiments yielded promising results with high cleavage activities of the two TALEN pairs within the GFP encoding sequence both in a restriction enzyme-like digestion and with a genomically integrated copy of the *gfp* gene in U-2 OS-GFP<sup>+</sup> cells. The higher activity of TALEN 7 and 8 in the previous two experiments led to the decision to continue the project with this pair. A prerequisite for the following experiments was a constant and high expression of the TALENs in the entire cell population. This setting is very much in contrast to TALENs applied during genomic engineering, which requires a transient expression, only. Since TALEN-mediated cytotoxic effects could not be excluded at this point, we decided to establish a cell line with a conditional expression system, which confers the expression of TALEN 7 and TALEN 8 upon addition of doxycycline (dox).

### 5.2.1 First induction and expression analysis

The conditional expression system utilizes the oriP-based expression plasmid p5252, which constitutively expresses the surface marker CD2. The expression cassettes TALEN 7:IRES:Cerulean and TALEN 8 are under the control of the dox-inducible bidirectional promoter (Fig. 5.3 A). HEK293 cells were stably transfected with this plasmid and selected by adding 500 ng/ml puromycin to the cell culture medium. The established cell line, referred to as 5252, was induced by the addition of 100 ng/ml dox and the cells were analyzed for the expression of Cerulean. Since the Cerulean encoding sequence is part of a bicistronic transcript and fused to the sequence coding for TALEN 7 it is as well under control of the bidirectional dox-inducible promoter. Therefore, Cerulean is a surrogate marker for the expression of TALEN 7 and TALEN 8. 24 hours after adding dox 49 % of the cells turned Cerulean-positive (Fig. 5.3 B).



**Fig. 5.3 Generation of two cell lines for the conditional expression of TALEN 7, TALEN 8, and Cerulean or mCherry**

HEK293 cells were stably transfected with two oriP-based expression plasmids. Both plasmids express the surface marker CD2 constitutively and harbor a bidirectional doxycycline-inducible minimal CMV promoter conferring conditional expression of nucleases and marker genes.

**A.** HEK293 cells stably transfected with p5252 led to the cell line 5252 expressing dox dependent TALEN 7, TALEN 8 and Cerulean. As indicated the left ORF codes for TALEN 7 fused to a NLS and a HA tag and the phenotypical marker Cerulean linked via an IRES element. The right ORF codes for TALEN 8 fused to a NLS and a Flag tag. The HA and Flag tags confer independent detection of TALEN 7 and 8 expression levels by Western blot immunostaining.

**B.** HEK293 cells stably transfected with the plasmid p5252 were compared regarding Cerulean expression by flow cytometry with and without dox. The addition of dox led to 49 % of Cerulean-positive cells (blue line) whereas Cerulean was not expressed in the absence of dox (black line).

**C.** Parental HEK293 cells and the 5252 cell line without and with dox were analyzed by Western blot immunostaining assessing the expression of TALEN 7 and 8. In the absence of dox the cells showed no signals but the addition of dox resulted in the appearance of a distinct band indicative of the expression of TALEN 7 (top panel) and TALEN 8 (middle panels). The staining with an  $\alpha$ -tubulin antibody (bottom panel) demonstrates the equal loading of the different protein extracts. The asterisk points to an unspecific signal.

**D.** The cell line 5257 was generated by stably transfecting HEK293 cells with the plasmid p5257. Addition of dox leads to the expression of mCherry (left ORF) and Luciferase (right ORF); the expression of the Luciferase gene has no implications in the experiments discussed in the course of this PhD work.

**E.** HEK293 cells stably transfected with p5257 were analyzed by flow cytometry to evaluate expression of mCherry. After the addition of dox, 56 % of the cells turned mCherry-positive (red line) whereas no positive cells were observed when dox was omitted (black line).

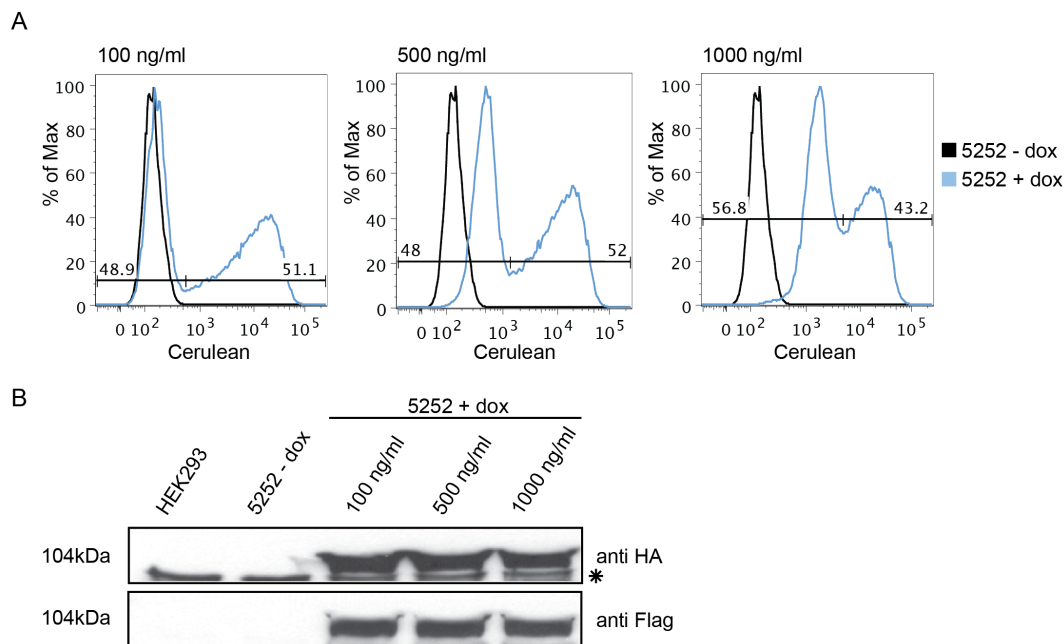
dox= doxycycline, NLS= nuclear localization signal



TALEN 7 and TALEN 8 were fused to a HA tag and a Flag tag, respectively. This allowed the independent determination of their expression levels via Western blot analysis. Upon the addition of dox both TALENs were expressed whereas no signal was detected when dox was omitted (Fig. 5.3 C). Additionally, a second cell line was established, which served as a negative control in the following experiments. It utilizes the oriP-based expression plasmid p5257, which confers the expression of mCherry and Luciferase upon induction with dox. mCherry was used as a marker for the induction rate whereas the expression of the Luciferase gene has no further implications in the course of this PhD work (Fig. 5.3 D). HEK293 cells were stably transfected with the plasmid and selected by adding 500 ng/ml puromycin to the cell culture medium. The generated cell line, referred to as 5257, was induced by the addition of 100 ng/ml dox and after 24 hours the expression of mCherry was assessed by flow cytometry. Induction with dox resulted in 56 % of mCherry positive cells (Fig. 5.3 E).

### 5.2.2 Comparison of different doxycycline concentrations

The 5252 cell line was generated to establish highly responsive cells, which homogeneously express the TALENs in the entire population upon induction. The induction of 100 ng/ml dox led to 49 % of Cerulean-positive cells, only. I consequently asked whether induction with higher dox concentrations might increase the percentage of Cerulean-positive cells and the expression levels of TALEN 7 and TALEN 8. For this purpose, cells were induced with 100 ng/ml, 500 ng/ml and 1000 ng/ml dox. After 24 hours the fraction of Cerulean-positive cells was assessed by flow cytometry and expression levels of TALEN 7 and TALEN 8 were analyzed by Western blot analysis. Adding different dox concentrations did not alter the fractions of Cerulean-positive cells nor the rate of induction. At higher dox concentrations an increasing fluorescence was observed reflecting the increased overall fluorescence of the aromatic compound dox (Fig. 5.4 A). Western blot analysis revealed that neither the expression level of TALEN 7 nor TALEN 8 was any different with altering dox concentrations (Fig. 5.4 B). Both experiments did not indicate a dose-dependent effect of dox concentration on the rate of induction or expression levels within the range of dox concentration tested.



**Fig. 5.4 Different dox concentrations do not alter the induction rate and protein expression levels of cells stably transfected with p5252.**

HEK293 cells stably transfected with the oriP-based expression vector p5252 (Fig. 5.3) were induced with different dox concentrations. After 24 hours, the induction rates and protein expression levels were assessed by flow cytometry and Western blot immunostaining, respectively.

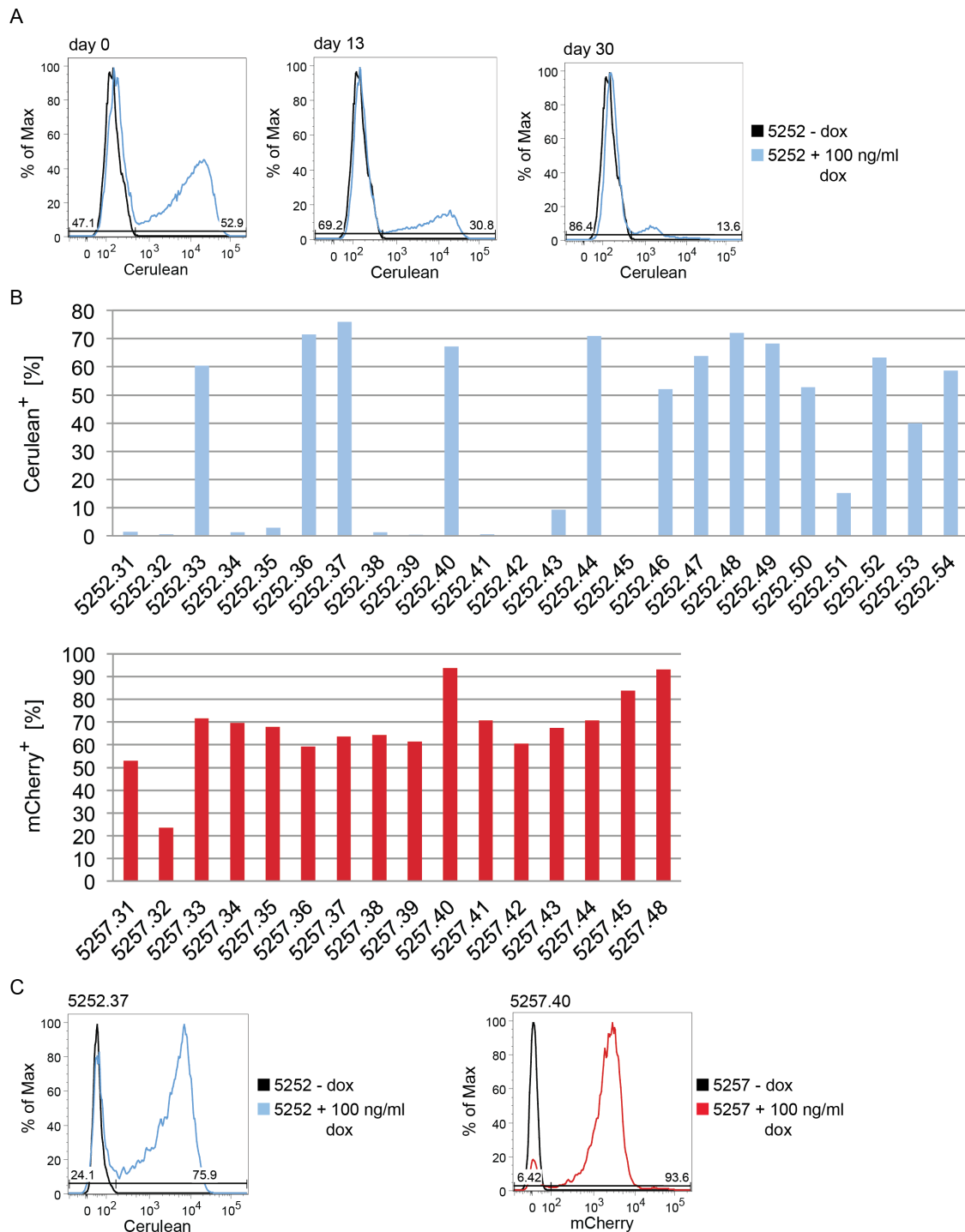
**A.** 5252 cells were analyzed in the absence (black line) or presence of dox (blue line). Addition of 100 ng/ml induced 51 % of the cells (left panel). Higher concentrations of dox (middle and right panel) showed no change in the rate of induction but led to a right shift of the whole population indicative of increased dox-mediated fluorescence of the cells.

**B.** Expression levels of TALEN 7 (α-HA tag) and TALEN 8 (α-Flag tag) were analyzed by Western blot immunostaining as a function of different dox concentrations. No signal was present in HEK293 or uninduced 5252 cells (first and second samples). The addition of 100 ng/ml, 500 ng/ml and 1000 ng/ml dox resulted in equal expression levels of TALEN 7 and TALEN 8 (third, fourth and fifth rows) indicating that the steady state expression levels of the nucleases are not affected by different dox concentrations. The asterisk points to an unspecific signal.

dox= doxycycline

### 5.2.3 Single cell clones and their dox inductivity

The induction of stably transfected and selected 5252 and 5257 cells resulted in about 50 % Cerulean-positive and mCherry-positive cells, only. Using higher dox concentrations did not increase the induction rate. This result was unexpected and therefore I analyzed the situation in more detail. 5252 cells were induced by adding dox and the fraction of Cerulean-positive cells was assessed by flow cytometry early after plasmid selection (day 0), 13 days and 30 days later.



**Fig. 5.5 HEK293 cell lines stably transfected with p5252 or p5257 indicate clonal variation of dox-mediated marker gene expression**

Clonal cell lines were established and analyzed by flow cytometry for their dox-mediated inducibility assessing the fraction of Cerulean- (5252) or mCherry-positive cells (5257).

**A.** The induction rate of bulk transfected 5252 cells was analyzed early after transfection and plasmid selection (left panel), 13 days (middle panel) and 30 days later (right panel) by flow cytometry. Early after selection 50 % of the cells were inducible (left panel). The fraction of inducible cells constantly decreased to 30 % on day 13 (middle panel) and 13 % on day 30 (right panel).

**B.** To establish clonal cell lines, single cells were picked and expanded to generate cell clones. 24 individual clones of the parental 5252 cell line were picked (upper panel) and analyzed for their inducibility of Cerulean. Only eight cell clones showed a clearly higher inducibility than the parental cell line, the majority of the cell clones were only weakly inducible or not at all. 16 clonal cell lines were generated from the mCherry expressing 5257 cell line (lower panel). All of them showed responsiveness

to dox and twelve out of 16 showed a clearly higher inducibility than the previously established bulk cell line.

C. Selected clones, which showed the highest rate of induction upon dox addition were characterized further. Clone #37 obtained from the 5252 cell line showed an induction rate of 75 % of Cerulean-positive cells (left panel); 93 % of Clone #40 derived from the parental 5257 cell line were mCherry-positive upon induction (right panel).

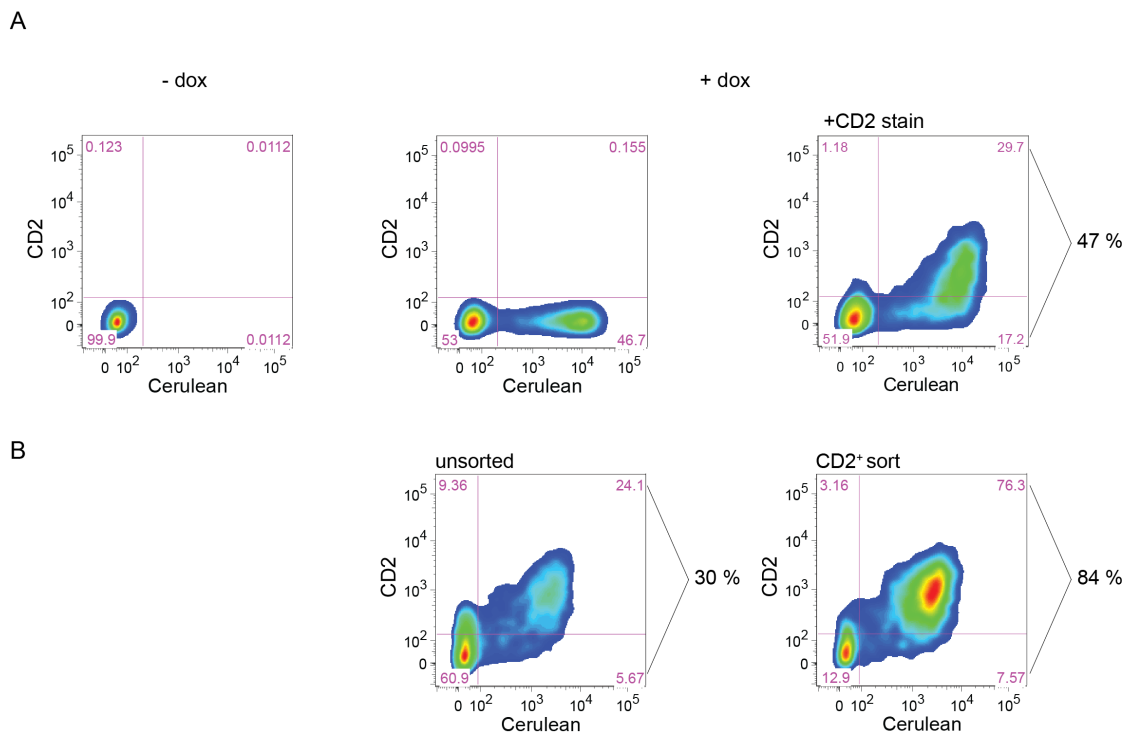
dox= doxycycline

The experiment showed that the dox responsiveness of the 5252 cell line was drastically decreasing (Fig. 5.5 A). These findings led to the decision to establish clonal cell lines derived from the 5252 and 5257 bulk populations in order to generate cell lines, which show high and consistent induction rates over a longer period of time. Of the TALEN expressing 5252 bulk population 24 clones were picked and analyzed further. Eight clones showed a clearly higher induction rate than the parental 5252 cell line whereas half of the clonal cell lines showed no or very weak inducibility, only (Fig. 5.5 B). 16 clones were generated from the mCherry expressing cell line 5257. Twelve of the 16 clones showed a clearly better inducibility than the parental cell line (Fig. 5.5 B). Together, my experiment suggests an interference of TALENs encoded by the plasmid p5252 and their dox-mediated gene expression. The clones, which showed the most convincing inducibility were 5252.37 and 5257.40 with 75 % Cerulean- and 93 % mCherry-positive cells, respectively (Fig. 5.5 C).

#### **5.2.4 Correlation between CD2 expression and induction rate**

The established clonal cell lines 5252.37 and 5257.40 achieved an initially higher rate of induction but this gain of function was not stable over time (data not shown). The question arose if this problem might originate from epigenetic modifications of the oriP plasmid leading to a repression of dox-mediated gene expression over time. To address this issue the population of 5252 cells was analyzed for the expression of the surface marker CD2, which is constitutively expressed by the SV40 early promoter and the expression of Cerulean by adding dox (Fig. 5.3 A). 5252 cells in the absence of dox showed no Cerulean expression. Upon addition of dox about half of the cells turned Cerulean-positive. Co-staining of the CD2 surface marker revealed that the majority of the Cerulean-positive cells expressed CD2, whereas the uninducible 5252 cells were almost entirely CD2-negative (Fig. 5.6 A). This finding led to the assumption that CD2

expression is a surrogate marker of dox inducibility. To evaluate this assumption dox was added to the population of 5252 cells, which led to 30 % Cerulean positive cells (Fig. 5.6 B). Subsequently, cells were enriched by magnetic cell sorting for CD2 expression. After magnetic cell sorting 84 % of the enriched cells expressed Cerulean, i.e. were responsive to dox (Fig. 5.6 B). Consequently, I used this correlation in a two-step protocol. In the first step I sorted the cells for CD2 expression, in the second step I added dox in order to achieve a better inducibility of the cells in most of the following experiments.



**Fig. 5.6 Dox-dependent expression of Cerulean correlates with constitutive expression of CD2**

HEK293 cells stably transfected with the oriP-based p5252 expression plasmid encoding CD2, Cerulean, TALEN 7, and TALEN 8 were analyzed by FACS. The CD2 gene is constitutively expressed from the SV40 early promoter. Cerulean, TALEN 7 and TALEN 8 genes are under the control of a dox inducible bidirectional promoter (Fig. 5.3). Dox inducibility of Cerulean was analyzed in relation to the expression levels of the surface molecule CD2.

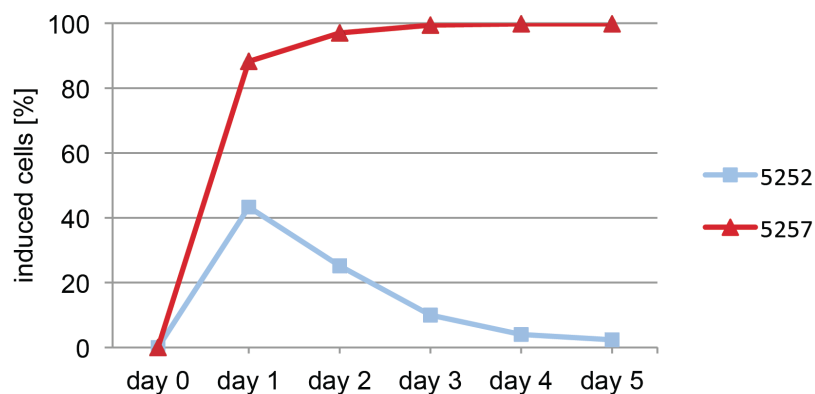
**A.** HEK293 cells carrying p5252 were analyzed for the expression of Cerulean in the absence (left panel) or 16 hours after the addition of dox (middle and right panel). Cerulean is not expressed without dox (left panel). About half of the cells (47 %) turned Cerulean-positive after adding dox (middle panel). Co-staining of the surface marker CD2 revealed that the majority of Cerulean-positive cells co-expressed CD2 (right panel top) suggesting that CD2 is a marker for dox inducibility.

**B.** The cells were enriched by magnetic cell sorting with antibody-coupled beads specific for CD2 (right panel) and compared to the parental cell population (left panel) after addition of dox for 16 hours. 30 % of the unsorted parental cell population turned Cerulean-positive (left panel) whereas 84 % of the sorted cells were dox responsive and expressed Cerulean (right panel).

dox=doxycycline

### 5.2.5 Kinetics of dox-induced gene expression

Another important parameter of the inducible system was the kinetics of gene expression after induction with dox. To address this aspect, 5252 and 5257 cells were induced with 100 ng/ml dox and the induction rates were assessed over time for five days (Fig. 5.7). The TALEN expressing cell line 5252 showed a maximum of dox-mediated gene expression on day one (40 % in this experiment) with a steady decrease on the following days to 2 % on day five. In contrast, the mCherry expressing cell line 5257 showed an 88 % rate of induction on day one, which increased further to almost 100 % on day five (Fig. 5.7). The results indicated that expression of TALENs clearly interfered with their dox-dependent inducibility very much in contrast to the marker protein mCherry.



**Fig. 5.7 Expression kinetics of cells expressing TALEN 7, TALEN 8 and Cerulean (5252) or mCherry (5257) after addition of dox**

HEK293 cells stably transfected with oriP-based expression plasmids p5252 or p5257 (Fig. 5.3) were induced by adding 100 ng/ml dox. The fractions of Cerulean-positive (5252) or mCherry-positive cells (5257) were measured by flow cytometry to assess the induction rates over time within a period of five days.

Upon induction 5252 cells reached 43 % of Cerulean-positive cells on day 1 post induction. Thereafter the fraction of Cerulean expressing cells decreased considerably dropping to 2 % on day 5 (blue line). In contrast, 88 % of the 5257 cells turned mCherry-positive on day one and increased to almost 100 % on day five (red line). The results indicate that the expression kinetics of TALENs differ considerably from the expression of a phenotypic marker protein.

dox= doxycycline

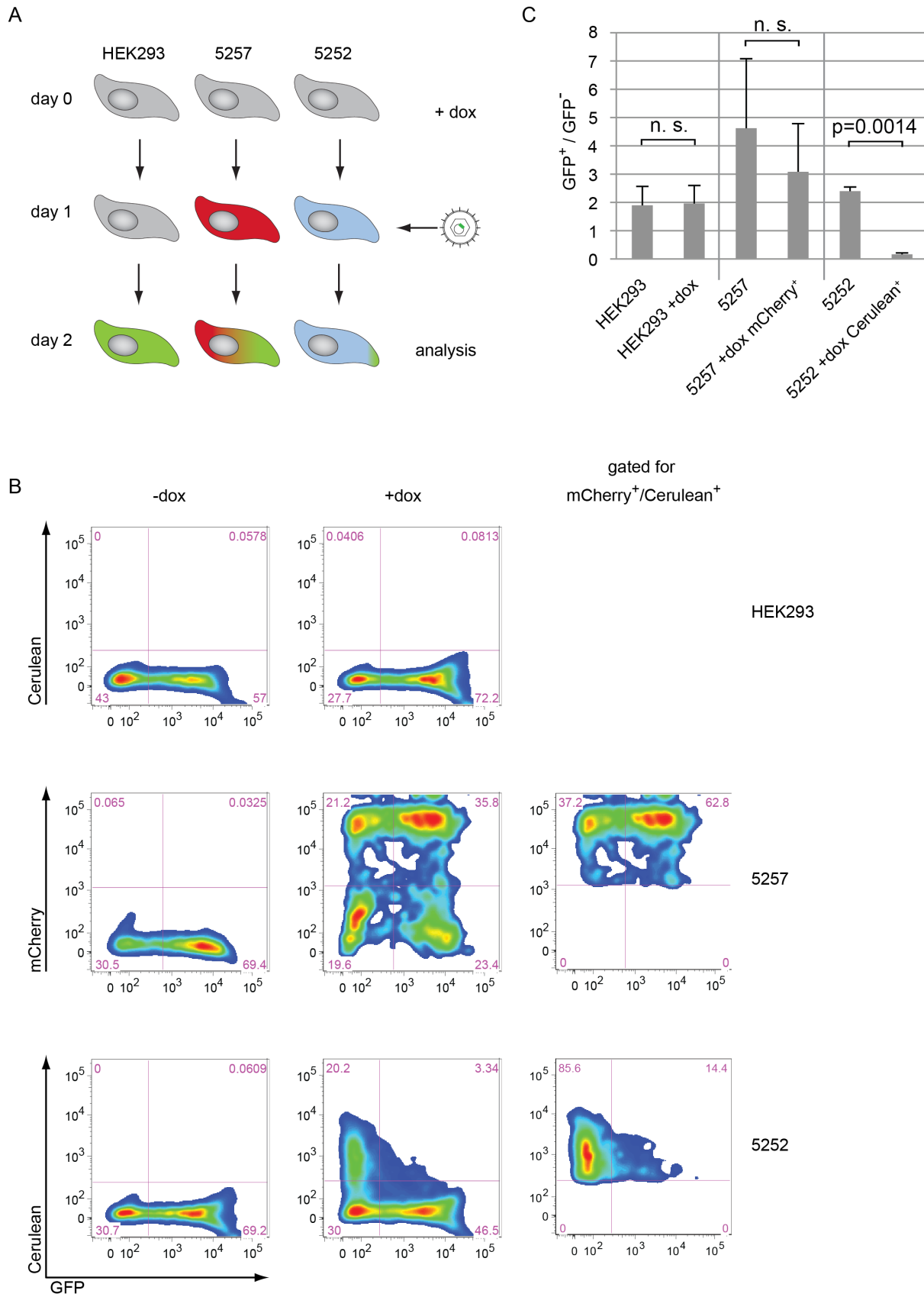
### 5.3 Infection of TALEN expressing cells with MHV-68-gfp

The previous experiments have shown that TALEN 7 and TALEN 8 successfully induce DNA double-strand breaks both *in vitro* (Fig. 5.1) and *in vivo* (Fig. 5.2). The next question was, if TALENs can also induce DSBs within the viral DNA upon infection with an appropriately engineered DNA virus. For this purpose the established 5252 and 5257 cell lines expressing TALEN 7, TALEN 8 and Cerulean or mCherry, respectively, were infected with MHV-68-gfp and subsequently analyzed by flow cytometry and qPCR. MHV-68-gfp is a BAC-derived version of the murine herpesvirus 68 harboring the *gfp* coding sequence.

#### 5.3.1 Knockout of viral GFP

The question if TALEN 7 and TALEN 8 induce a DSB within viral DNA was first addressed at the level of GFP protein expression. The three different cell lines HEK293, 5252 and 5257 were infected with the virus MHV-68-gfp. The virally encoded *gfp* sequence contains the cognate sequence motifs of TALEN 7 and 8. On day 0 dox was added to the three cell lines. In the 5252 cells dox induces TALEN 7, TALEN 8 and Cerulean expression, in the 5257 cells mCherry is induced and dox-treated parental HEK293 cells served as the negative control. On day one the cells were infected with MHV-68-gfp. 24 hours later on day two the cells were analyzed by flow cytometry. HEK293 cells were expected to turn GFP-positive upon infection. 5257 cells expressing mCherry were expected not to interfere with viral infection or GFP expression and thus should result in a mCherry-positive and GFP-positive population. The expression of TALEN 7 and 8 in 5252 cells was expected to result in the cleavage of the viral *gfp* sequence and hence a Cerulean-positive but predominantly GFP-negative population was expected in this experimental setting (Fig. 5.8 A).

On day two, the cells were analyzed for GFP expression (x-axis) and Cerulean or mCherry expression (y-axis) (Fig. 5.8 B). The analysis of HEK293 cells showed that dox had no impact on the fraction of GFP-positive cells after infection with MHV-68-gfp compared to HEK293 cells without dox indicating that the compound does not interfere



**Fig. 5.8 The expression of TALEN 7 and TALEN 8 reduces the expression of the MHV-68-gfp encoded *gfp* gene after infection**

Two derivatives of the parental HEK293 cell line stably transfected with oriP-derived plasmids were engineered, conferring dox-dependent expression of TALEN 7 and TALEN 8 and appropriate marker



genes. HEK293 cells as well as the two cell lines were infected with the murine virus MHV-68-gfp, which encompasses the coding sequence of the *gfp* gene. Upon addition of dox, the cell line 5252 expresses the proteins TALEN 7, TALEN 8 and Cerulean whereas the cell line 5257 expresses mCherry, only, as a control. The cell lines were analyzed by flow cytometry assessing the fractions of GFP-positive and GFP-negative cells as a function of TALEN 7, TALEN 8 expression after infection.

**A.** Experimental overview: On day 0 parental HEK293 cells and the two cell lines 5252 and 5257 were induced by adding dox. After 24 hours (day 1) cells, which were responsive to dox, expressed mCherry (5257) or TALEN 7, TALEN 8 and Cerulean (5252). On day one HEK293 cells and the two cell lines were infected with MHV-68-gfp, which harbors the *gfp* gene encompassing the target sequence of TALEN 7 and 8. On day two cells were analyzed by flow cytometry. 5257 cells expressing mCherry upon addition of dox, only, were expected to show no interference with GFP expression or viral infection and therefore an infection with MHV-68-gfp was expected to generate a mCherry-positive and a GFP-positive population. Cleavage of the viral *gfp* sequence after infection in 5252 cells was expected to result in a Cerulean-positive but predominantly GFP-negative population.

**B.** Cells were analyzed for the expression of Cerulean or mCherry (y-axis) and GFP (x-axis) by flow cytometry. In the first row, parental HEK293 cells were analyzed and served as a negative control. A comparable fraction of cells (57 and 72 %) turned GFP-positive after infection with MHV-68-gfp as expected, demonstrating that the addition of dox does not interfere with viral infection or GFP expression. In the second row, 5257 cells were analyzed in the absence (left panel) or presence of dox inducing the expression of mCherry (middle panel). Without dox 69 % of the cells turned GFP-positive after infection (left panel) comparable to parental HEK293 cells (57 %; top row). After adding dox 57 % (21 % + 36 %) of the cells turned mCherry-positive (middle panel). Of the mCherry-positive population 63 % were GFP-positive after infection (right panel). This result indicates that mCherry expression does not interfere with viral infection, because uninduced 5257, induced 5257 or parental HEK293 cells in the absence or presence of dox showed comparable populations of GFP-positive cells (69 %, 59 %, 72 %, and 57 %, respectively). The third row shows the results of the 5252 cell line after infection with MHV-68-gfp without dox (left panel) and after adding dox leading to the expression of TALEN 7, TALEN 8 and Cerulean (middle panel). In the absence of dox viral infection resulted in 69 % GFP-positive cells (left panel) similar to parental HEK293 and 5257 cells above. In the experiment shown addition of dox resulted in a total of 23 % Cerulean-positive cells, which express TALEN 7 and TALEN 8 (20 % + 3 %; middle panel). Of the Cerulean-positive population, only 14 % turned GFP-positive after viral infection; a clear reduction (right panel) compared to HEK293 or 5257 cells (+/- dox) or the uninduced 5252 cell line. The majority of Cerulean-positive cells (86 %) that expressed TALEN 7 and 8 did not turn GFP-positive indicative of a TALEN-mediated cleavage of the virally encoded *gfp* gene.

**C.** The diagram summarizes the results of the three different cell lines infected with the same dose of MHV-68-gfp virus without and with dox in three independent experiments. The results are displayed as the ratios of GFP-positive to GFP-negative cells. HEK293 cells show the same ratio of GFP-positive to GFP-negative cells in the absence or presence of dox. About 70 % of the HEK293 cells turn GFP-positive resulting in a ratio of about two. Similarly, the 5257 cell line, which expresses mCherry upon dox induction showed comparable ratios of GFP-positive to GFP-negative cell fractions in the absence or presence of dox. For unknown reasons, 5257 cells consistently showed a higher fraction of GFP-positive cells as compared to parental HEK293 cell line infected with the same dose and multiplicity of infection with MHV-68-gfp virus. Infection of the uninduced 5252 cell line resulted in the same ratio of GFP-positive to GFP-negative cells as the parental HEK293 cells. Upon dox-mediated expression of TALEN 7 and TALEN 8 the ratio of GFP-positive versus GFP-negative cells was massively reduced indicating a successful cleavage of the virally encoded *gfp* gene.

dox= doxycycline

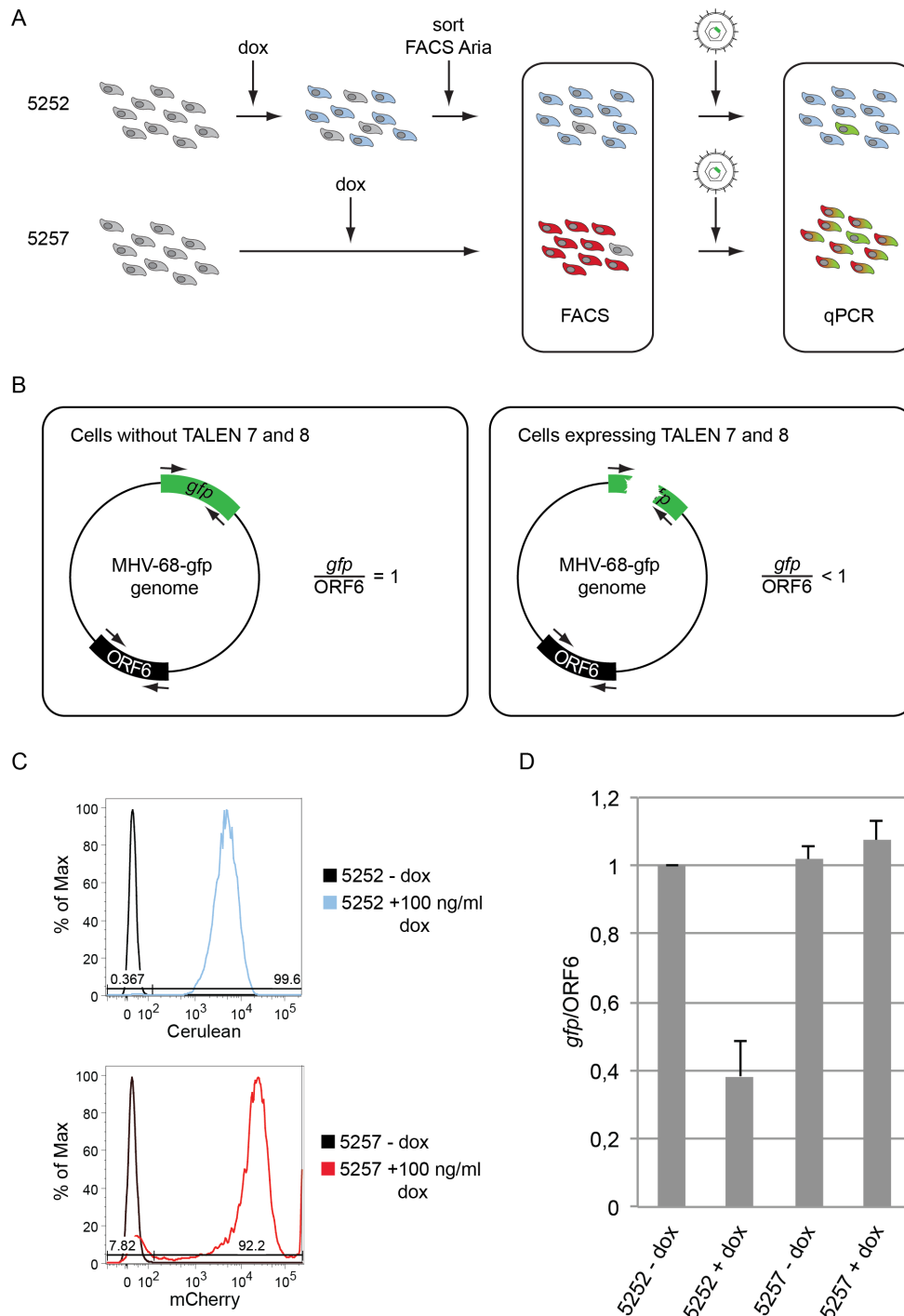
with viral infection or GFP expression (Fig. 5.8 B, top row). Similarly, the infection of 5257 cells in the absence of dox resulted in 69 % of GFP-positive cells. The analysis of infected 5257 cells in the presence of dox, revealed 62 % of GFP-positive cells in the mCherry-positive fraction indicating no interference of mCherry expression with viral infection or GFP expression. The infection of uninduced 5252 cells resulted in a fraction of 69 % of GFP-positive cells comparable to HEK293 and 5252 cells with or without dox. The analysis of induced 5252 cells, revealed only 14 % of GFP-positive cells within the

Cerulean-positive population. This expected result most likely documents the cleavage of the viral *gfp* sequence by TALEN 7 and 8 and their interference with GFP expression (Fig. 5.8 B). With the result of three independent experiments I computed the ratio of GFP-positive to GFP-negative cells shown in Fig. 5.8 C. I found no significant difference in HEK293 (+/- dox) and 5257 cells (+/- dox). In contrast, addition of dox and the concomitant expression of TALEN 7 and TALEN 8 resulted in a significant decrease of the ratio of GFP-positive to GFP-negative cells in line with TALEN-induced DSB within the virally encoded *gfp* sequence (Fig. 5.8 C).

### 5.3.2 Analysis of TALEN-mediated cleavage of the viral *gfp* sequence by qPCR

The previous experiment showed that TALEN 7 and 8 interfere with the expression of the virally encoded *gfp* gene. In the next experiment I re-evaluated this finding and employed a PCR strategy to document the cleavage of the *gfp* locus by the TALENs as predicted. Towards this end, the two cell lines 5252 and 5257 were induced by adding dox. 5252 cells were sorted for Cerulean-positive cells to reach at least 90 % of induced cells. 5252 and 5257 cells were analyzed for their rate of induction by flow cytometry and subsequently infected with the same dose and multiplicity of infection of MHV-68-*gfp* virus stock. 24 hours post infection total DNA was extracted and analyzed by qPCR (Fig. 5.9 A). qPCR analysis was performed with primers amplifying either the *gfp* sequence or the essential viral gene ORF6 to allow normalization of the data. In the absence of nucleases the ratio of the intact *gfp* and ORF6 genes should be close to one. The expression of TALEN 7 and 8 was expected to induce a double-strand break at a predicted site within the viral *gfp* sequence (Fig. 4.2). Cleavage of the *gfp* gene abrogates PCR amplification, therefore qPCR should reveal a ratio of *gfp* versus ORF6 that is considerably lower than one (Fig. 5.9 B). Prior to infection the induction rate was assessed by flow cytometry since a highly induced population is necessary to record clear effects. In each experiment, induction rates exceeded 90 % (Fig. 5.9 C). For the calculation of the three different experiments the qPCR results of uninduced 5252 cells were set to one and the qPCR results of the other samples were normalized accordingly. The induction of TALEN expression in 5252 cells resulted in a *gfp*/ORF6 ratio of

approximately 0.4 indicative of a TALEN-mediated double-strand break within the viral *gfp* sequence (Fig. 5.9 D). The qPCR results of uninduced and induced 5257 cells revealed *gfp*/ORF6 ratios of about one as expected (Fig. 5.9 D).



**Fig. 5.9 Cleavage of the viral *gfp* sequence after infection of TALEN-expressing cells**

Two cell lines, which express TALEN 7, TALEN 8 and Cerulean (transfected with p5252) or mCherry (transfected with p5257) were infected with the *gfp* encoding virus MHV-68-gfp. 24 hours after infection the total cell DNA was isolated and the integrity of *gfp* sequences was assessed by qPCR.

**A.** Experimental overview: Addition of dox-induced protein expression of the dox-dependent genes in the 5252 and 5257 cell lines. Induced 5252 cells were physically sorted to reach a purity of at least 90 %. This step was not necessary for 5257 cells. The rate of induction was assessed by flow cytometry. Cells were

infected with MHV-68-gfp and after 24 hours total DNA was extracted from both cell lines to be analyzed by qPCR.

**B.** The scheme shows two possible outcomes of the experiment depicted in panel A. The left panel shows the status of the viral genome after infection of cells that do not express TALEN 7 and 8; the *gfp* sequence stays intact. ORF6 codes for an essential viral gene and is used as a normalization control for quantitative PCR. Since both genes are unique in the viral genome the relative quantification of *gfp* to ORF6 results in a ratio of one. If TALEN 7 and 8 expressing cells were infected with MHV-68-gfp a cleavage of the *gfp* sequence is expected to occur. A cleaved copy of *gfp* cannot serve as a template in qPCR analysis, therefore, the ratio of *gfp* versus ORF6 was expected to drop below one (right panel).

**C.** The induction rate of both cell lines was assessed prior to infection with MHV-68-gfp. The 5252 cell line was induced after cell sorting exceeding 99 % Cerulean-positive cells (upper panel). The induction rate of 5257 cells was higher than 92 % (lower panel). Both cell lines showed no detectable expression of the marker genes in the absence of dox.

**D.** Summary of experimental results. The ratios of *gfp* versus ORF6 in uninduced 5252 cells (first bar) were set to one. The results of the other samples were normalized to this sample. The expression of TALEN 7 and 8 resulted in a *gfp* to ORF6 ratio of 0.38 indicating that the *gfp* coding sequence within the viral genome was cleaved in the majority of the cells (second bar). When 5257 cells were infected with MHV-68-gfp in the absence or presence of dox (third and fourth bar) the ratio of *gfp* to ORF6 was about one demonstrating that neither dox nor the induced expression of the marker gene interferes with the qPCR results.

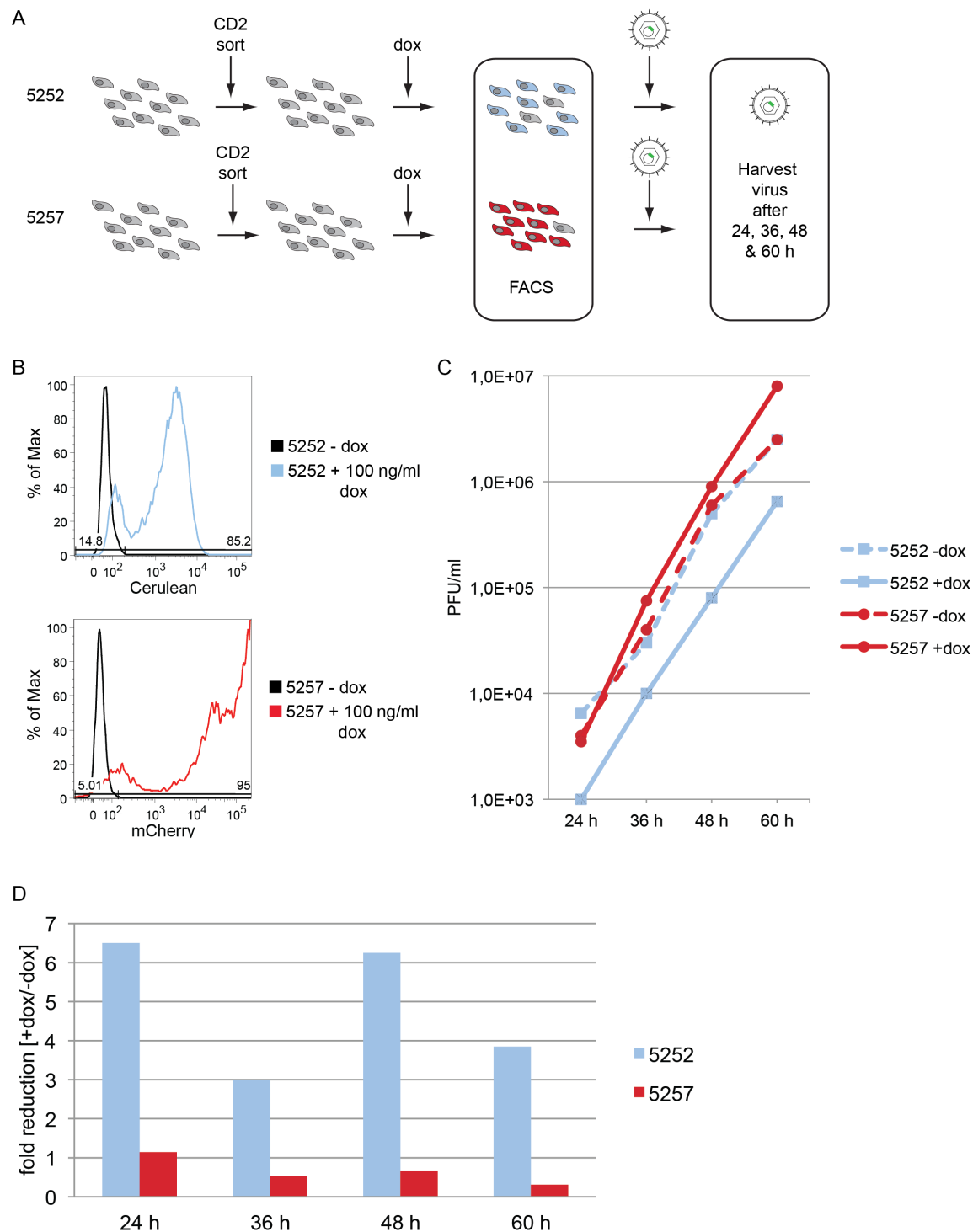
dox= doxycycline, ORF= open reading frame

## 5.4 Interference of TALEN expression with viral *de novo* synthesis

The previous two experiments revealed that TALEN 7 and 8 cleave their cognate sequence within the viral *gfp* gene of MHV-68-gfp. A double-strand break within the viral genome was expected to block viral *de novo* synthesis because the viral genome is fragmented and cannot form circles upon infection, which are a prerequisite for viral DNA replication in the lytic phase of infection. To address this issue plaque assays were performed, which allow the quantification of newly synthesized virus in a biologically relevant assay.

### 5.4.1 MHV-68-gfp *de novo* synthesis in TALEN or mCherry expressing cells

The concentration of newly synthesized virions were compared in the supernatants of the TALEN 7, TALEN 8 and Cerulean expressing cell line 5252 and the mCherry expressing cell line 5257. Both cell lines were sorted for the surface marker CD2, which



**Fig. 5.10 Expression of TALEN 7 and 8 in 5252 cells reduces *de novo* synthesis of MHV-68-gfp after viral infection**

Standard viral plaque assays were performed to compare *de novo* synthesis of MHV-68-gfp after infection of HEK293 cells expressing TALEN 7, TALEN 8 and Cerulean (5252) or control cells expressing mCherry (5257), only.

**A.** Experimental overview: Initially, both cell lines were sorted for CD2 surface expression to maximize the fraction of dox-inducible cells as in Fig. 5.6. Dox was added to the CD2 enriched cells. Twelve hours later the induction rate of Cerulean was assessed by flow cytometry and cells were infected with MHV-68-gfp. Cell supernatants containing newly synthesized virions were harvested 24, 36, 48, and 60 hours after infection and quantified by plaque assays.

**B.** Flow cytometry analysis showed induction rates of 85 % for 5252 cells (upper panel) and 95 % for 5257 cells (lower panel) 12 hours after adding dox. In the absence of dox no induction was observed in both cell lines (upper and lower panels).

**C.** Quantification of infectious virions in the supernatants was performed by plaque assays 24, 36, 48, and 60 hours after infection of 5252 and 5257 cells. The addition of dox to 5257 cells (solid red line) did not alter the amount of virus particles as measured by plaque forming units (PFU) per ml supernatant compared to uninduced 5257 cells (dotted red line). In the absence of dox, 5252 cells (dotted blue line) supported virus synthesis at a level comparable to 5257 cells (red lines). The addition of dox to 5252 cells resulted in a clear reduction of viral *de novo* synthesis at all time points (solid blue line).

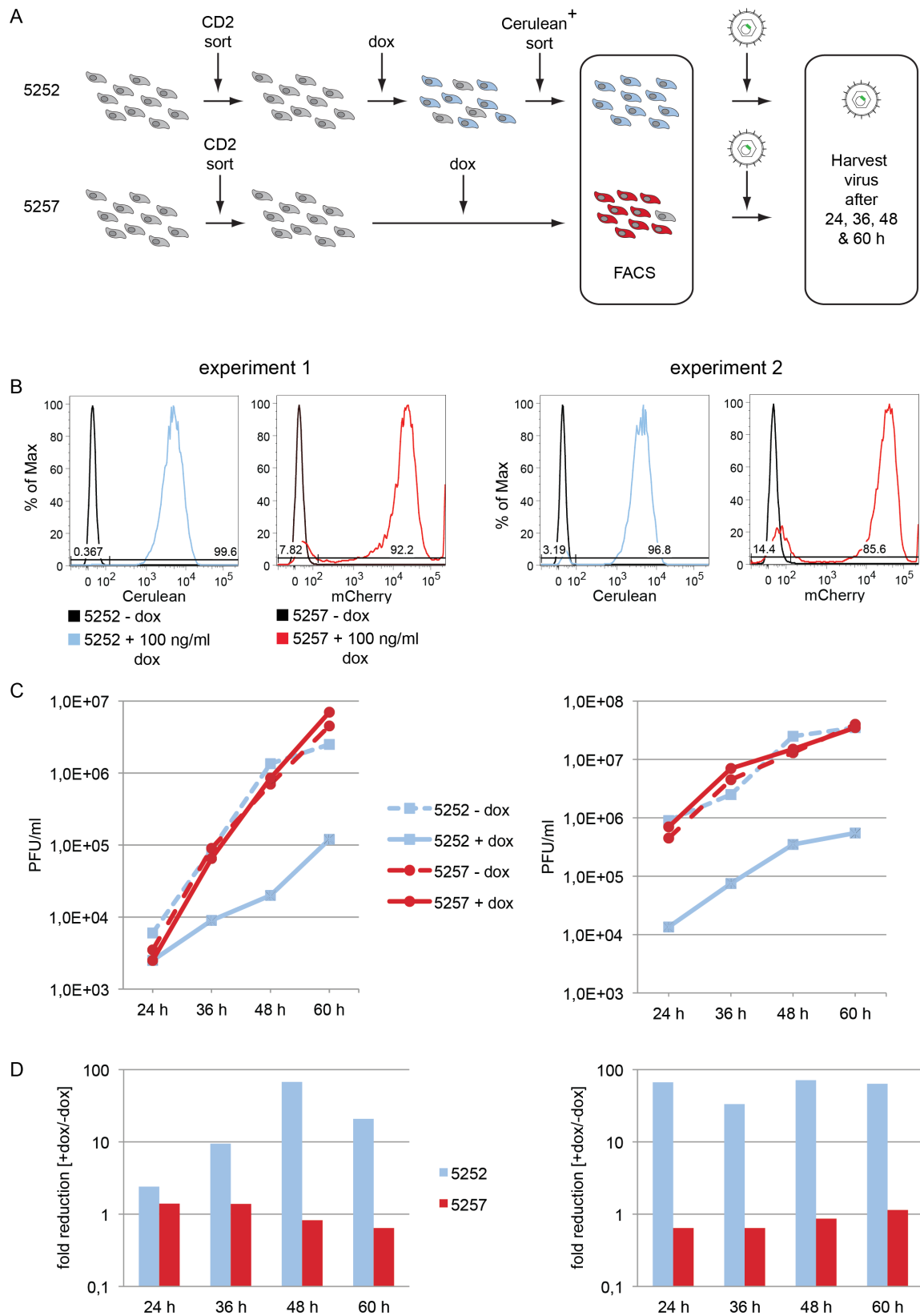
**D.** The differences of infectious virions were calculated by computing the ratio of virions in the supernatants of uninduced cells to induced cells. The ratio in mCherry expressing 5257 cells is at all time points close to one (red bars) whereas the TALEN 7 and 8 expressing 5252 cells show a clear reduction at all time points with a maximum at 24 and 48 hours (blue bars) indicating that cleavage of the viral genome hampers viral *de novo* synthesis.

dox= doxycycline, PFU= plaque forming units

maximizes the inducibility of the cell populations as shown in Figure 5.6. 24 hours after the addition of dox cells were analyzed for the expression of Cerulean (5252) or mCherry (5257) and subsequently infected with MVH-68-gfp. The cell supernatants, which contain the newly synthesized virions were harvested 24, 36, 48, and 60 hours after viral infection and quantified by plaque assays (Fig. 5.10 A). The rates of dox-mediated induction in this experiment were 85 % and 95 % in 5252 and 5257 cells, respectively (Fig. 5.10 B). The quantification of harvested viral samples revealed that the addition of dox to 5257 cells did not interfere with viral replication as expected. In contrast, dox-mediated TALEN expression in 5252 cells resulted in a clear reduction of newly synthesized virus at all time points compared to uninduced 5252 cells (Fig. 5.10 C). The ratios of virions in the supernatants of induced versus uninduced 5257 cells were close to one at all time points indicative of a comparable degree of viral *de novo* synthesis. In the TALEN expressing 5252 cells viral *de novo* synthesis was impaired at all time points. At 24 and 48 hours after infection virus synthesis was maximally reduced by a factor of 6. In 5252 cells, only 85 % of the cells were induced (Fig. 5.10 B), which limits the calculated theoretical reduction to a factor of 6.7. The results indicate that a TALEN-mediated DNA double-strand break within the viral genome in principal can prevent viral *de novo* synthesis (Fig. 5.10 D).

#### **5.4.2 MHV-68-gfp *de novo* synthesis in highly induced TALEN or mCherry expressing cells**

The first plaque assay performed in the previous experiment showed that TALEN expression reduces viral replication. The fraction of TALEN-positive cells were limited in



**Fig. 5.11 Induced expression of TALEN 7 and 8 in 5252 cells reduces *de novo* synthesis of MHV-68-gfp after viral infection**

Plaque assays were performed with the two different cell lines 5252 and 5257 infected with MHV-68-gfp. The viral *gfp* sequence harbors the target sequence of TALEN 7 and 8 resulting in a cleavage of the viral genome (Fig. 5.8 and Fig. 5.9). 5252 cells express TALEN 7, TALEN 8 and Cerulean upon induction with

dox. 5257 cells serve as a control and express mCherry, only. Viral *de novo* synthesis was compared in both cell lines without and with dox-induced protein expression.

**A.** Experimental overview: 5252 and 5257 cell lines were sorted for CD2 surface expression, which correlates with a high rate of induction (Fig. 5.6). Twelve hours after dox-mediated induction 5252 cells were sorted for Cerulean-positive cells enriching highly induced cells. In both cell lines induction rates were assessed by flow cytometry. Subsequently, cells were infected with the same dose and multiplicity of infections with MHV-68-gfp virus. Newly synthesized virus was harvested 24, 36, 48, and 60 hours post infection and quantified by plaque assays.

**B.** Analysis of both cell lines by flow cytometry twelve hours after addition of dox demonstrated the high rates of induced protein expression prior to infection with MHV-68-gfp. The experiment was performed twice. The induction rates of the 5252 and 5257 cell lines in the two experiments are shown in panels 1 to 4. Cell sorting of induced 5252 cells resulted in 99 % Cerulean-positive population (panel 1). 92 % of the induced 5257 cells were mCherry positive (panel 2). The induction rates in the second experiment are shown in panel 3 and 4. Expression rates of Cerulean and mCherry in 5252 and 5257 cells revealed 96 % and 85 %, respectively.

**C.** 24, 36, 48, and 60 hours after infection of the two cell populations with MHV-68-gfp viral samples were harvested from infected induced and uninduced cells. Quantification of infectious virions (PFUs) was performed by plaque assays. The left panel shows the results of the first experiment. In 5257 cells the addition of dox and concomitant expression of mCherry did not interfere with *de novo* synthesis of MHV-68-gfp (solid and dashed red lines). In 5252 cells in the absence of dox virus synthesis (dashed blue line) was comparable to uninduced and induced 5257 cells (dashed and solid red lines). The induced expression of TALEN 7 and 8 in 5252 cells massively reduced viral *de novo* synthesis indicating a TALEN-mediated inhibition of viral replication. These results were confirmed in a second experiment (right panel).

**D.** The reduction of virus concentration was calculated by computing the ratio of virions in uninduced versus induced cells. In the first experiment the induction of TALEN 7 and 8 expression (5252) resulted in a reduction of PFU by a factor of 67 48 hours after infection (blue bars) whereas the induction of mCherry expression did not change viral *de novo* synthesis indicated by a ratio of approximately one (red bars). In the second experiment (right panel) 5252 cells showed a strong reduction of viral synthesis at all time points which reached a factor of 71 48 hours after infection (blue bars). In mCherry expressing 5257 cells the factor was at all time points close to one (red bars) as expected from the control cells.

dox= doxycycline, PFU= plaque forming units

this pilot experiment (5.4.1) but the reduction of *de novo* synthesized virus was optimal.

I repeated the experiment with the cell lines 5252 and 5257 with one modification. The cell lines were sorted prior to dox-mediated induction for the expression of the surface marker CD2, which correlates with a high inducibility as shown in Figure 5.6. Subsequently, cells were induced by adding dox and the 5252 cells were additionally sorted for Cerulean-positive cells to enrich the TALEN expressing cells. Induction rates were assessed by flow cytometry prior to infection with MHV-68-gfp. Newly synthesized virus was harvested 24, 36, 48, and 60 hours post infection and quantified by plaque assays (Fig. 5.11 A). The induced expression of TALENs in 5252 cells is a crucial parameter as described above and was assessed twelve hours post induction by flow cytometry. In both experiments, induction rates of TALEN expressing 5252 cells exceeded 95 % (Fig. 5.11 B). The induction rate of mCherry expressing 5257 cells was 92 % in the first and 85 % in the second experiment (Fig. 5.11 B). In both experiments uninduced and induced 5257 cells supported viral replication equally, again indicating that the expression of mCherry does not interfere with *de novo* synthesis of the virus.



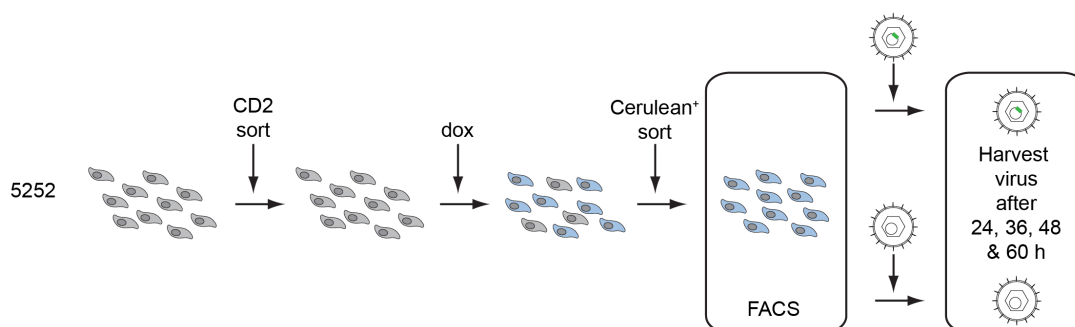
Uninduced 5252 cells showed comparable results to 5257 cells +/- dox. Upon addition of dox and concomitant expression of TALEN 7 and TALEN 8 viral *de novo* synthesis was drastically reduced in two independent experiments confirming the TALEN-mediated inhibition of viral replication already observed in the previous experiment (compare Fig. 5.10 C and Fig. 5.11 C). The ratios of virions in the supernatant of induced mCherry expressing 5257 cells versus uninduced cells was about one as expected at all time points. In contrast, TALEN expressing 5252 cells supported viral replication at much lower levels. Viral *de novo* synthesis was reduced by a factor of 71 in one of the two experiments (experiment 2) 48 h post infection (Fig 5.11 D).

#### **5.4.3 A comparison of MHV-68 and MHV-68-gfp *de novo* synthesis in TALEN expressing HEK293 cells**

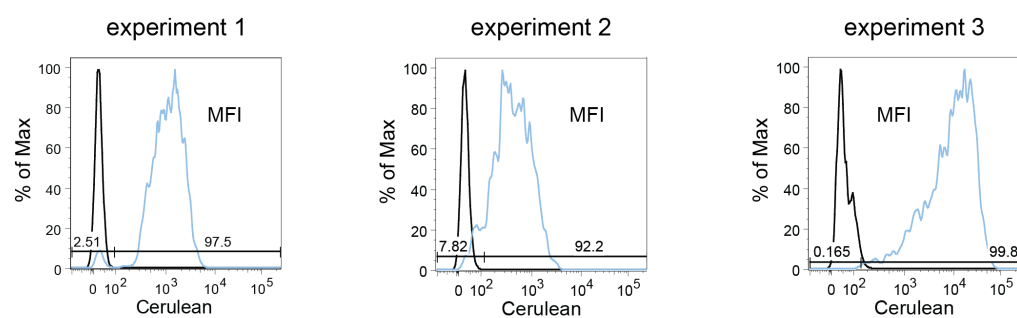
Infection of TALEN expressing 5252 cells with MHV-68-gfp showed a strong reduction of viral *de novo* synthesis compared to uninduced 5252 cells and the mCherry expressing control cell line 5257. This result suggested a TALEN-mediated block of viral *de novo* synthesis. This effect could be unspecific and mediated by a general genotoxic effect of the nucleases or due to a specific cleavage of the viral genome by TALEN 7 and TALEN 8. To address this pivotal question, TALEN expressing 5252 cells were infected with MHV-68-gfp, similar to previous experiments, and MHV-68 as a control. MHV-68-gfp harbors the *gfp* coding sequence and thus the cognate sequence of TALEN 7 and TALEN 8. In contrast, MHV-68 lacks the *gfp* coding sequence; consequently it should not be cleaved by TALEN 7 and 8 and therefore replicate unaffected in induced 5252 cells. In order to achieve a highly induced population 5252 cells were sorted for the surface expression of CD2, which correlates with a high inducibility (Fig. 5.6). Subsequently, cells were induced by adding dox and twelve hours later again sorted for Cerulean-positive cells in a second step. The fraction of induced cells was assessed by flow cytometry and infected with the same dose and multiplicity of infection of MHV-68-gfp or MHV-68 virus stocks (Fig 5.12 A, B). After 24, 36, 48, and 60 hours *de novo* synthesized virus was harvested and quantified by plaque assays (Fig. 5.12 A). The experiment was performed three times. In each experiment the induction rate exceeded

90 % (Fig. 5.12 B). *De novo* synthesis of MHV-68-gfp was reduced in TALEN 7 and 8 expressing 5252 cells as expected.

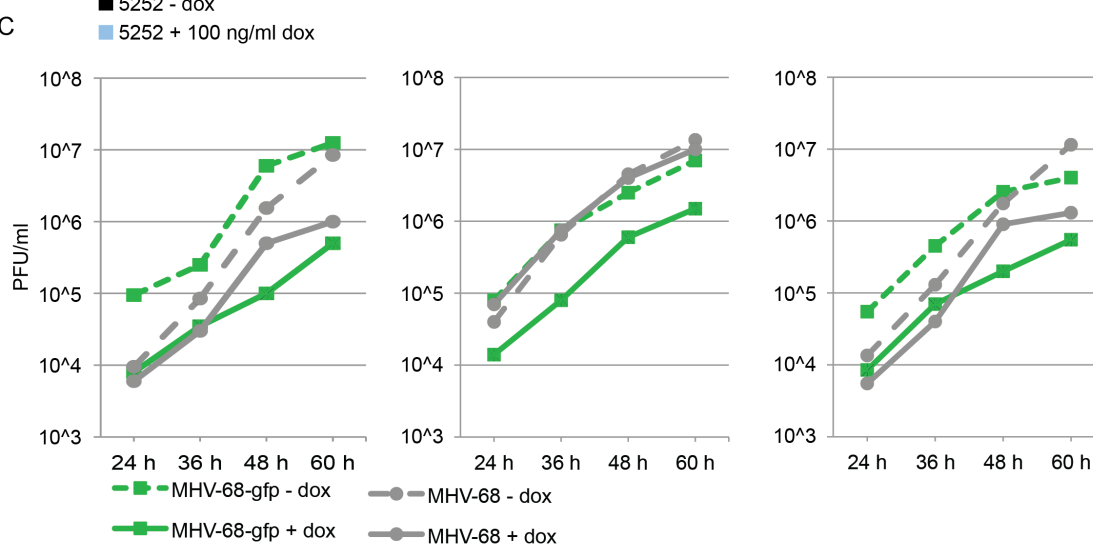
A



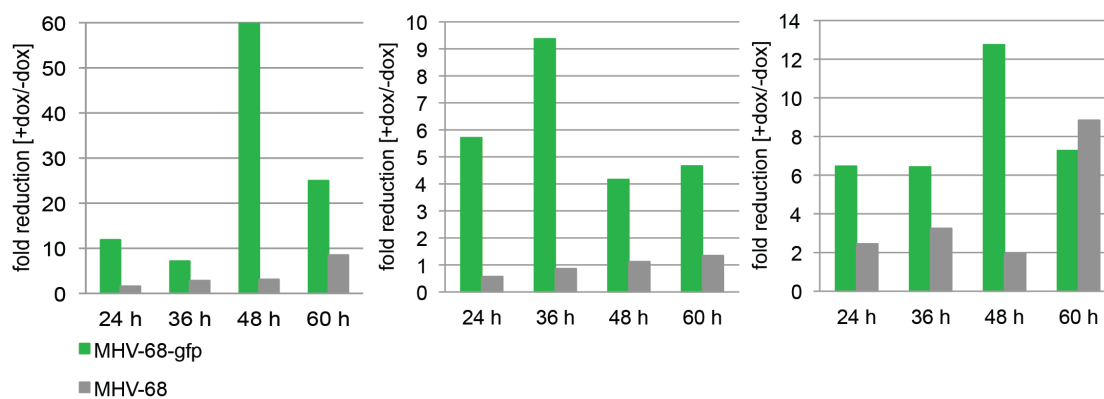
B



C



D



**Fig. 5.12 MHV-68-gfp but not MHV-68 *de novo* synthesis is reduced in TALEN 7 and 8 expressing 5252 cells**

Cells stably transfected with the oriP-derived expression plasmid p5252 were infected with two different viruses. MHV-68-gfp contains the *gfp* sequence and provides the target sequence for TALEN 7 and 8-mediated cleavage. Induction of TALEN 7 and 8 caused the impaired viral synthesis of MHV-68-gfp (Fig. 5.10 and Fig. 5.11) in 5252 cells. The control virus MHV-68 does not contain the TALEN 7 and 8 target sequence. Replication of both viruses in uninduced and induced 5252 cells was compared in plaque assays.

**A.** Experimental overview: 5252 cells were sorted for CD2 to maximize the rate of induction (Fig. 5.6) prior to the addition of dox. Twelve hours post induction Cerulean-positive 5252 cells were enriched by cell sorting to obtain a highly induced population as assessed by flow cytometry. Subsequently, cells were infected with the same dose and multiplicity of infections with MHV-68-gfp or MHV-68 virus stocks. Newly synthesized virus was harvested after 24, 36, 48, and 60 hours and quantified by plaque assays.

**B.** The induction rates of Cerulean after cell sorting were assessed by flow cytometry. The results of three different experiments are shown in the three panels. In all experiments the induction rate of Cerulean was higher than 90 %.

**C.** In all experiments viral *de novo* synthesis of the MHV-68-gfp virus was reduced at all time points in induced cells (solid green line) compared to uninduced cells (dashed green line) as expected. An infection with the *gfp*-negative virus MHV-68 showed no or only minor differences between uninduced and induced cells at the time points 24, 36 and 48 hours (dashed and solid grey lines). An exception was the 60 hours time point in experiment 1 and 3. A reduced synthesis of MHV-68 virus, lacking the target sequences of TALEN 7 and 8 are observed (exp 1 and exp 3; left and right panels) suggesting that the expression of the endonucleases could impair virus synthesis globally at late time points in certain experiments.

**D.** The reduction of newly synthesized virus in induced versus uninduced cells was calculated for each time point. In all experiments MHV-68-gfp *de novo* synthesis was impaired upon addition of dox (green bars) when the cells expressed TALEN 7 and 8. The extent of reduction varied between a factor of 60 in the first experiment 48 hours post infection (left panel) and a factor of 9 in the second experiment 36 hours post infection (middle panel). Reduction of virus synthesis (Fig. 5.12 D) correlated with the different rates of gene induction (Fig. 5.12 B). In cells infected with MHV-68 the ratio was about one at all time points except at 60 h in experiment 1 and 3. Here, viral replication was reduced at a factor of 9.

dox= doxycycline, PFU= plaque forming units

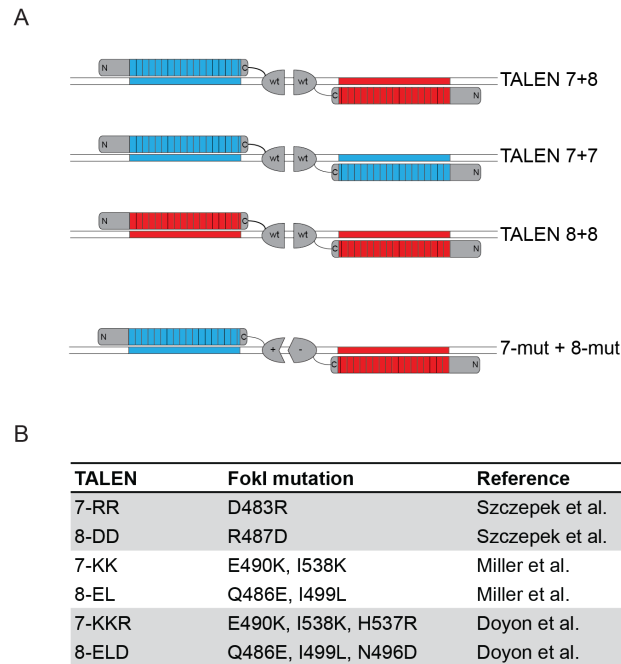
In contrast, infections with the *gfp*-negative virus MHV-68 revealed no or only minor differences between uninduced and induced cells except at the time point 60 hours post infection in experiment 1 and 3 (Fig. 5.12 C). Here, MHV-68 *de novo* synthesis was repressed up to a factor of 9 indicating an unspecific global TALEN-mediated effect (Fig. 5.12 C). The computed ratios of virus concentrations in the supernatants of induced versus uninduced 5252 cells revealed that replication of MHV-68-gfp was hampered in each of the three experiments at all time points. In the second experiment the reduction of MHV-68-gfp replication was less pronounced as compared to experiments 1 and 3 (Fig. 5.12 C). MHV-68 replication was unaffected in experiment 2. Together my findings correlated with a comparable low expression of TALEN 7 and 8 in experiment 2 (Fig. 5.12 B, D).

## 5.5 Generation of TALENs with mutated FokI domains

Infection of TALEN 7 and 8 expressing 5252 cells with MHV-68-gfp revealed a clear reduction of viral *de novo* synthesis, but the previous experiment also showed that TALENs can have a global negative effect on the replication of the MHV-68 control virus lacking the cognate sequence of TALEN 7 and 8 (Fig. 5.12). This finding suggests unspecific effects of the synthetic endonucleases on the host cell in general or the viral infection. The genotoxic effect of synthetic ZFNs is known and several groups have generated mutated FokI domains, which favor heterodimerization and show reduced genotoxic side-effects. In the framework of this thesis different mutated FokI domains were generated to further optimize the TALEN-based antiviral system.

### 5.5.1 Overview of TALENs with mutated FokI domains

A TALEN consists of the TAL effector DNA binding domain fused to the nuclease domain of the FokI restriction enzyme. In order to become catalytically active the FokI domains and therefore the TALENs have to dimerize. Since the FokI domains form homodimers the desired combination of TALEN 7 and TALEN 8 is not the only possible combination. A homodimerization of TALEN 7 molecules is also possible as well as a homodimer of TALEN 8 (Fig. 5.13 A). Several groups have shown that single amino acid substitutions within the FokI domains are sufficient to favor heterodimerization (Fig. 5.13 A). In the course of this thesis three pairs of heterodimerizing FokI domains were generated and compared regarding their cleavage activity and genotoxicity. The Figure 5.13 B gives an overview of the three different pairs and the relevant publications.



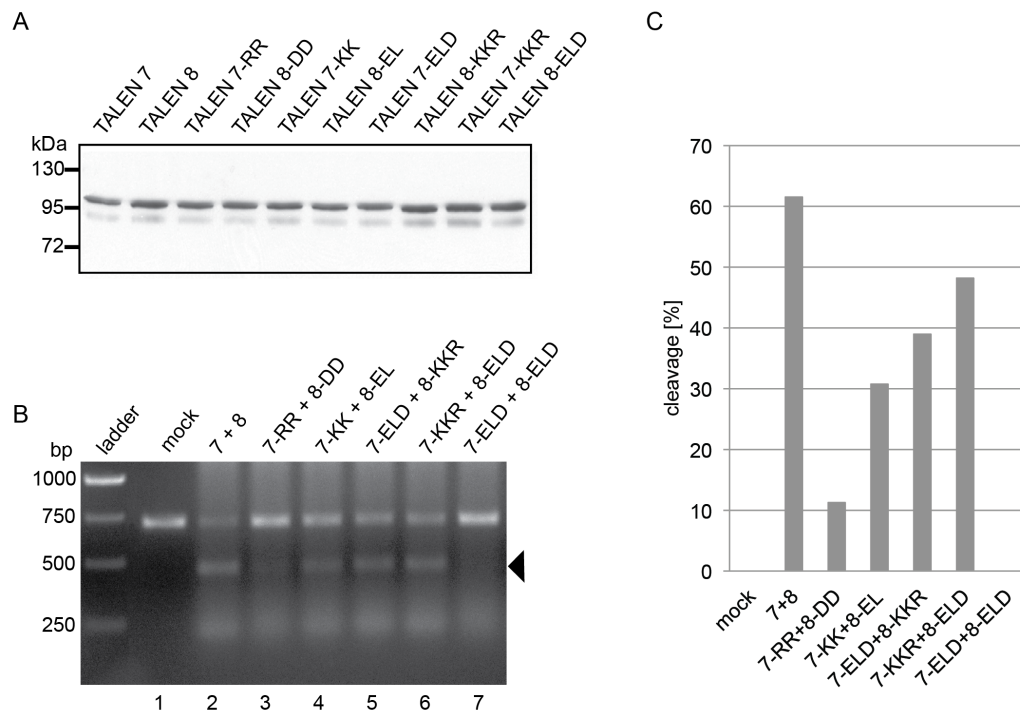
**Fig. 5.13 Single amino acid substitutions in the FokI domains prevent homodimerization and increase TALEN specificity**

**A.** The FokI nuclease domain forms homodimers. As a consequence, dimerization of TALEN 7 and 8 (upper panel) is not the only possible combination. Homodimerization of two TALEN 7 molecules (second panel) is possible as is homodimerization of TALEN 8 (third panel). Several groups have shown different mutations in the FokI cleavage and dimerization domain, which force heterodimerization (fourth panel).

**B.** Mutations within the FokI nuclease and dimerization domain do not only affect dimerization characteristics but also cleavage activity of the nucleases. Therefore, three different TALEN pairs with mutated FokI domains were generated and compared in this work. The table lists the generated TALEN pairs in the first column. The second column shows the introduced amino acid substitutions; the third column lists the original publications, with the individual FokI mutations.

### 5.5.2 *In vitro* cleavage assay comparing TALENs with mutated FokI domains

TALEN 7 and 8 with wild type and mutated FokI domains, shown in Figure 5.13, were compared in the *in vitro* cleavage assay, which was performed as shown in Figure 5.1. Western blot analysis showed comparable expression levels of all endonucleases (Fig. 5.14 A). The results of the *in vitro* cleavage assay revealed that TALEN 7 and 8 with wild type FokI domains have the highest cleavage activity (Fig. 5.14 B). The TALEN pair 7-RR and 8-DD showed a considerably lower cleavage activity. TALEN 7-KK and TALEN 8-EL performed better but not as well as TALEN 7-KKR and TALEN 8-ELD, which showed the highest activity. A reciprocal exchange of the FokI domains resulting in TALEN 7-ELD and TALEN 8-KKR showed a slightly reduced activity suggesting that the combination of



**Fig. 5.14 TALEN 7-KKR and TALEN 8-ELD is the pair with the highest *in vitro* cleavage activity**

The cleavage activities of the generated TALEN pairs with mutated FokI domains were compared in an *in vitro* cleavage assay. The experiment was performed as described in Fig. 5.1.

**A.** Western blot immunostaining indicates comparable expression levels of the different pairs of TALEN 7 and TALEN 8 (107 kDa).

**B.** Results of the *in vitro* cleavage assay were analysed on an ethidium bromide stained agarose gel. Lane 1 shows the *gfp* input DNA as a mock control. TALEN 7 and 8 with wild type FokI domains show the highest cleavage activity as the input DNA is cleaved almost completely (lane 2). TALEN 7-RR and TALEN 8-DD have hardly any cleavage activity since only a faint band appears (lane 3). TALEN 7-KK and TALEN 8-EL show a slightly higher activity compared to TALEN 7-RR and 8-DD (lane 4). TALEN 7-KKR and TALEN 8-ELD show the highest cleavage among the different mutated TALEN pairs (lane 6) but perform only slightly better than TALEN 7-ELD and TALEN 8-KKR (lane 5) in which the FokI domains were reciprocally exchanged. The addition of TALEN 7-ELD and TALEN 8-ELD shows that no cleavage occurs when a dimer formation of the mutated FokI domains is prevented as expected (lane 7). The arrowhead indicates the expected TALEN-mediated cleavage products.

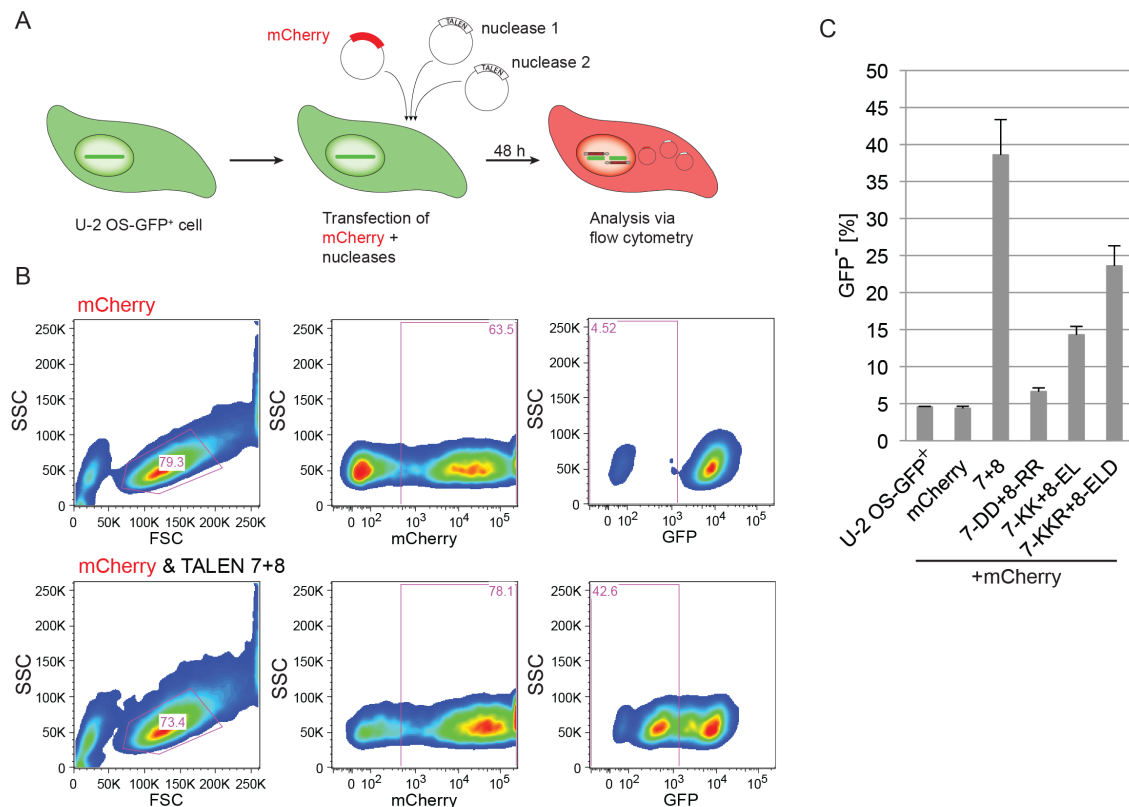
**C.** A quantification of the cleavage activities of the different TALEN pairs is shown. The first column is the mock control when no nucleases were added. TALEN 7 and 8 show the highest activity with more than 60 % of input DNA cleaved (column 2). The combination of RR and DD mutated TALENs almost completely abolishes cleavage activity (column 3). Addition of TALEN 7-KK and TALEN 8-EL shows reasonable cleavage activity (column 4). The FokI mutants ELD and KKR have the highest activity among the mutated FokI domains tested (column 5 and 6). The combination of TALEN 7-KKR and TALEN 8-ELD performed slightly better (column 6) The combination of TALEN 7-ELD and TALEN 8-ELD showed no cleavage as expected as the two proteins cannot form dimers.

TALENs and different mutated FokI domains has no major impact on the cleavage activity as expected. TALEN 7-ELD and TALEN 8-ELD are not able to form a dimer and consequently no cleavage was observed (Fig. 5.14 B). A quantification of the assay revealed that TALEN 7 and 8 cleaved more than 60 % of the input DNA (Fig. 5.14 C). TALEN pair 7-RR and 8-DD and TALEN pair 7-KK and 8-EL showed cleavage of about 11 % and 30 %, respectively. TALEN 7-KKR and 8-ELD showed the highest activity among TALENs with mutated FokI domains cleaving around 48 % of the template

whereas a reciprocal exchange of the FokI domains yielded around 39 % of cleaved product. The combination of TALEN 7-ELD and 8-ELD showed no cleavage since a dimerization of the FokI domains is not possible (Fig. 5.14 C).

### **5.5.3 *In vivo* cleavage assay comparing TALENs with mutated FokI domains**

The *in vivo* assay described in Figure 5.15 A was utilized to compare the cleavage activity of TALEN 7 and 8 equipped with different FokI domains. The assay was performed as described previously with U-2 OS-GFP<sup>+</sup> cells. After co-transfection of three plasmids the cells were first gated for survivors according to forward and sideward scatter criteria and to successfully transfected, mCherry-positive cells. Within this transfected population the proportion of GFP-negative and -positive cells was assessed. U-2 OS-GFP<sup>+</sup> cells transfected with mCherry showed 4 % of GFP-negative cells whereas co-transfection of TALEN 7 and 8 expressing plasmids led to 42 % of GFP-negative cells (Fig. 5.15 B). The summary of three independent experiments revealed no difference in the percentage of GFP-negative cells concerning untransfected cells or cells transfected with mCherry, only. Co-transfection with TALEN 7 and 8 yielded 38 % of GFP-negative cells. Co-transfection of TALEN 7-DD and 8-RR did not increase the percentage of GFP-negative cells (6%), which is close to the background level of untransfected cells or cells transfected with mCherry, only. TALEN 7-KK and 8-EL showed an intermediate phenotype with 14 % GFP-negative cells. Similar to the *in vitro* assay TALEN 7-KKR and 8-ELD reached the best result among TALENs with mutated FokI domains with 23 % of GFP-negative cells (Fig. 5.15 C).



**Fig. 5.15 TALEN 7-KKR and TALEN 8-ELD show the highest GFP knockout potential in U-2 OS-GFP<sup>+</sup> cells**

Cleavage activities of TALEN constructs with different FokI mutants were assessed after their transient expression in U-2 OS-GFP<sup>+</sup> cells, which carry a single chromosomal integrated copy of the targeted *gfp* gene sequence.

**A.** Experimental overview: GFP-positive U-2 OS cells (left panel) were co-transfected with three expression plasmids coding for (i) mCherry as a transfection control (ii) nuclease 1 and (iii) nuclease 2 (middle panel). After transfection, the cells were analyzed by flow cytometry after 48 hours. Cells were gated for mCherry expressing cells to restrict the subsequent analysis to transfected cells, only. In the population of mCherry-positive cells the proportion of cells that had lost GFP expression was assessed (right panel).

**B.** The analytical workflow is depicted. U-2 OS-GFP<sup>+</sup> cells were transfected with mCherry, only (upper row) or co-transfected with mCherry, TALEN 7 and TALEN 8 (lower row). In the left column the gates delineate the living cell populations, which were further analyzed in the middle panels. In the middle column cells were gated for the mCherry-positive populations to restrict analysis to transfected cells, only. In the right column the proportion of GFP-negative cells was assessed indicating the cleavage of the *gfp* gene. Cells co-transfected with constructs coding for TALEN 7 and TALEN 8 contained 42 % GFP-negative cells in contrast to cells transfected with mCherry, only.

**C.** The diagram summarizes the results of U-2 OS cells co-transfected with different pairs of TALEN 7 and TALEN 8 or mCherry, only. Cells transfected with the latter had a comparable amount of GFP-negative cells similar to untransfected U-2 OS cells (about 5 %). The co-transfection of TALEN 7 and 8 equipped with wild type FokI domains led to 38 % GFP-negative cells on average in three independent experiments. The mutant pairs TALEN 7-DD, TALEN 8-RR and TALEN 7-KK, TALEN 8-EL yielded only 6 % and 14 % GFP-negative cells, respectively. Co-transfection of TALEN 7-KKR and TALEN 8-ELD reached the best result (23 % GFP-negative cells) among TALEN pairs with mutated FokI domains.

FSC=forward scatter, SSC=sideward scatter

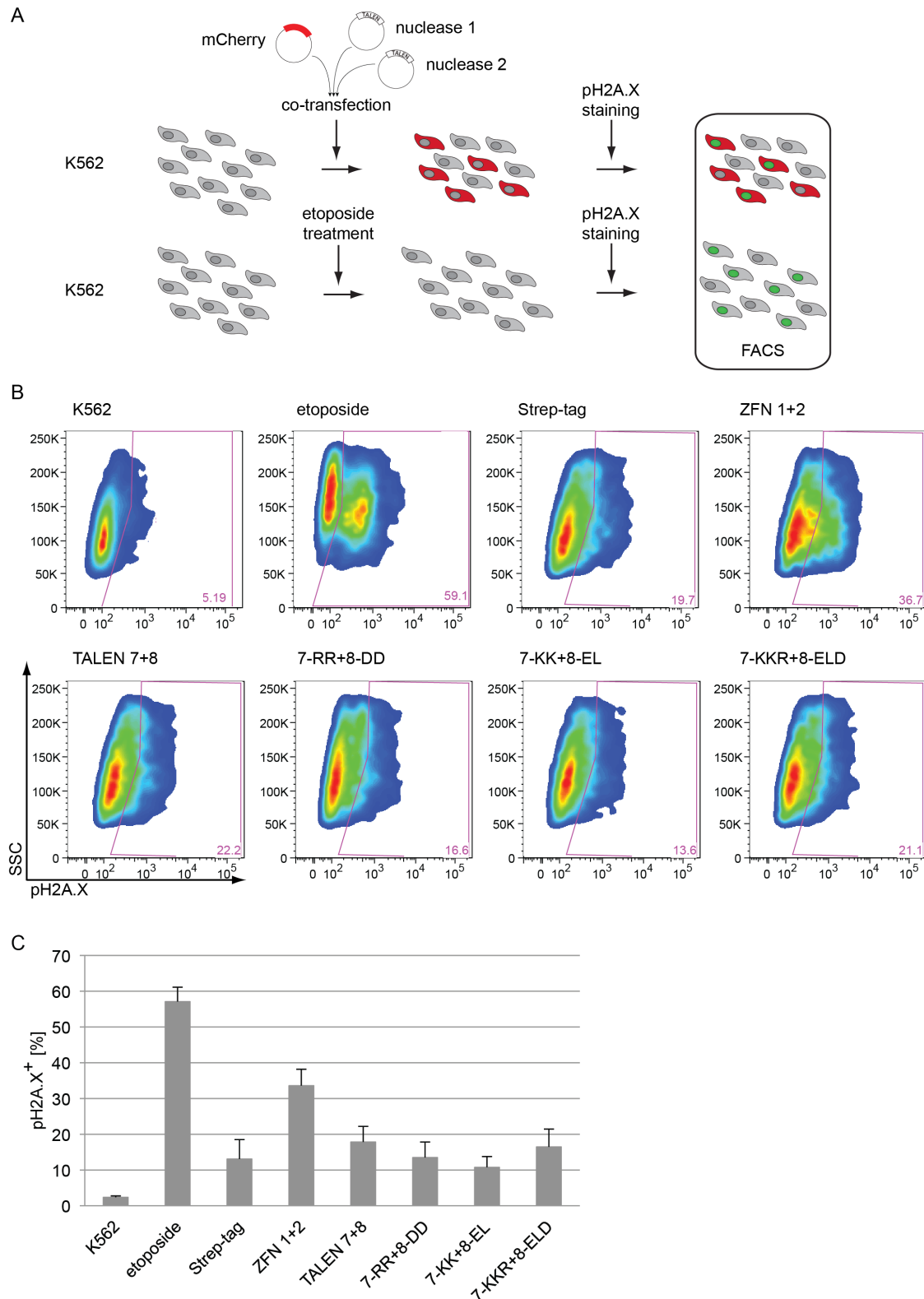


## 5.6 Analysis of TALEN-mediated genotoxicity

Plaque assays showed that TALENs can cleave viral DNA and thereby block viral *de novo* synthesis (Fig. 5.11) but they also indicated a TALEN-mediated cellular genotoxicity (Fig. 5.12). Because a TALEN-mediated genotoxicity is not compatible with a long-term expression of the nucleases, TALEN 7 and 8 were equipped with mutated FokI domains, which were expected to reduce unspecific cleavage and genotoxicity. In chapter 5.5 I compared the cleavage activities of TALEN 7, TALEN 8 and their mutant derivatives. In the following experiments I assessed the genotoxicity of the different TALENs.

### 5.6.1 DNA damage response in cells expressing different TALEN pairs

To assess the genotoxic potentials of TALEN 7, TALEN 8 and their derivatives with mutated FokI domains K562 cells were co-transfected with expression plasmids encoding either a TALEN or a ZFN or mCherry as a transfection control. 48 hours after transfection cells were analyzed for phosphorylated histone H2A.X, which is a marker for DNA double-strand breaks by intracellular flow cytometry. K562 cellular DNA does not contain the TALENs target sequence, therefore enhanced pH2A.X signals result from unspecific cleavage events in cellular DNA. Etoposide is used in cancer therapy as a chemotherapeutic drug. It is cytotoxic and known to induce DNA strand breaks. Here, etoposide is used as a positive control (Fig 5.16 A). 5 % of untransfected K562 cells were found to be pH2A.X-positive, but after etoposide treatment 59 % of the cells turned positive (Fig. 5.16 B). As a negative control, cells were co-transfected with mCherry and an empty expression vector (Strep-tag). Transfection resulted in a rather high background of 19 % pH2A.X positive cells (Fig. 5.16 B, top row 3<sup>rd</sup> panel). The combination of ZFN 1 and ZFN 2 delivered a signal of 36 % positive cells indicating a high degree of DNA damage by unspecific cleavage. Co-transfection of wild type TALEN 7 and 8 or different derivatives with mutated FokI domains showed results of 22, 16, 13 and 21 % of pH2A.X positive cells (Fig. 5.16 B, bottom row). These values are close to the value of the negative control (19 %), i.e. cells transfected with the empty expression plasmid (Strep-tag), only (Fig. 5.16 B).



**Fig. 5.16 DNA damage response in K562 cells transfected with TALENs or ZFNs**

**A.** Experimental overview: K562 cells were co-transfected with three plasmids encoding (i) mCherry as a transfection control, (ii) nuclease 1 and (iii) nuclease 2. After 48 hours cells were stained for pH2A.X, which indicates DNA double-strand breaks. The DNA of K562 cells does not contain the target sites of the nucleases, consequently, they should not induce double-strand breaks in K562 DNA. Phosphorylation of H2A.X upon transient expression of synthetic endonucleases would indicate unspecific DNA cleavage. After intracellular staining for pH2A.X the cells were analyzed by flow cytometry. The cells were gated for mCherry-positive cells to restrict the analysis to successfully co-transfected cells (upper panel). As a

positive control, K562 cells were treated with etoposide. Etoposide induces DNA double-strand breaks and therefore delivers a strong DNA damage response signal (lower panel).

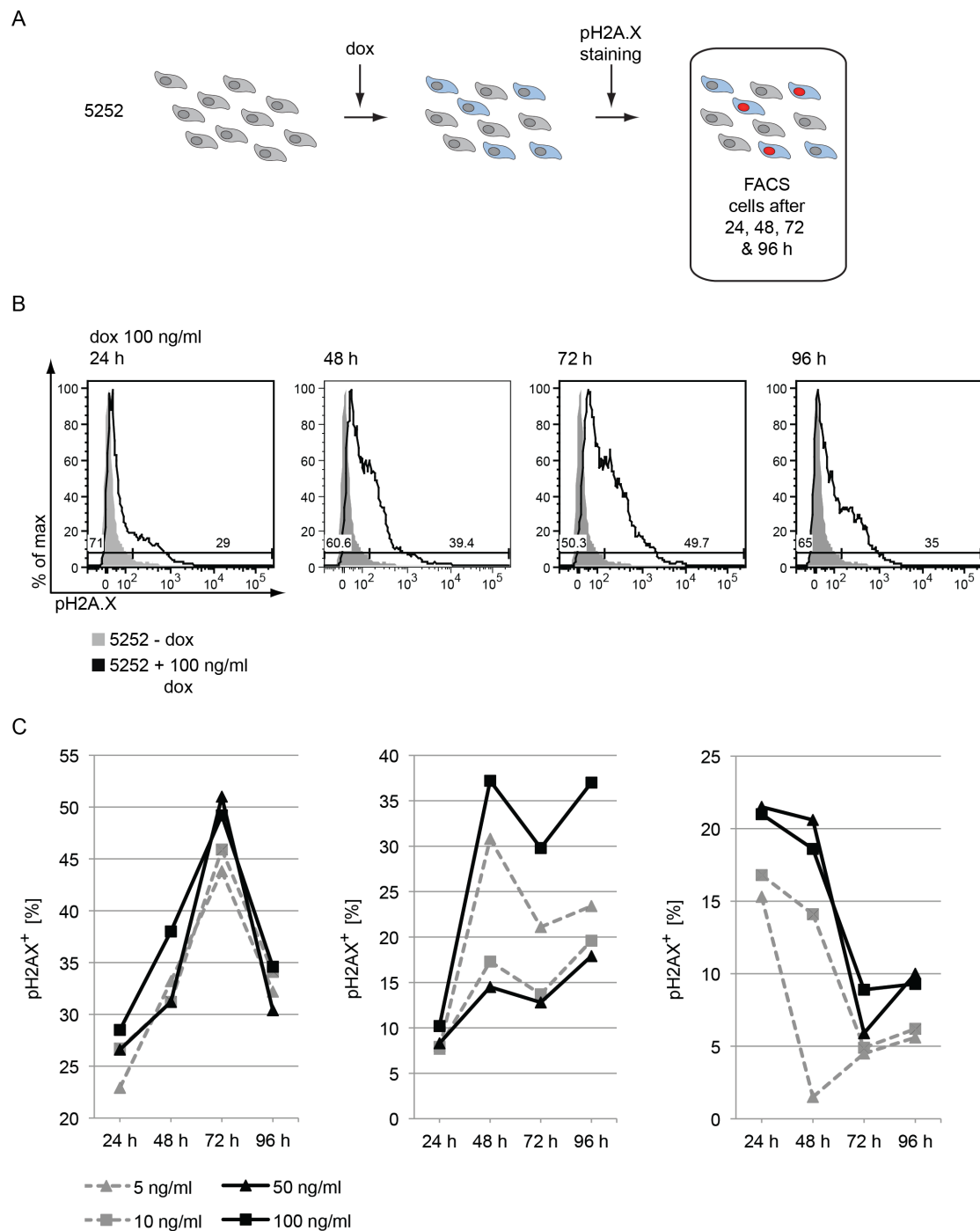
**B.** Results of a representative experiment are shown. Untransfected K562 cells showed 5 % of pH2A.X positive cells; treatment with etoposide led to 59 % of pH2A.X positive cells (upper row, panel 1 and 2). K562 cells transfected with the mCherry encoding plasmid and an empty expression plasmid (Strep-tag) were used as a negative control and showed 19 % of pH2A.X positive cells. The increase in pH2A.X positive cells suggests a high background in this assay (upper row, panel 3). Co-transfection of ZFN 1 and 2 with mCherry delivered a strong signal of 36 % pH2A.X positive cells indicating a high degree of DNA damage by ZFN 1 and 2 (upper row, panel 4). In the lower row K562 cells co-transfected with TALEN 7 and 8 with wild type FokI domains (lower row, panel 1) or FokI domains with different mutations (lower row, panel 2, 3 and 4) were compared. The TALEN transfected cells revealed 22, 16, 13 and 21 % pH2A.X positive cells comparable to cells transfected with the control plasmid (Strep-tag), only (upper row, panel 3).

**C.** The table summarizes the results of three different independent experiments. Etoposide (second bar) induced a high fraction of pH2A.X positive cells compared to untreated cells (first bar). Co-transfection with an empty expression vector (third bar, Strep-tag) caused a clear increase in pH2A.X positive cells compared to K562 cells (first bar). ZFN 1 and 2 induced pH2A.X in about 37 % of the transfected cells (fourth bar) indicating unspecific cleavage by the ZFN 1 and 2 pair. Co-transfections with the four different pairs of TALENs (fifth to eighth bar) showed signals in the range of the negative control (third bar) indicating a reduced genotoxicity compared to ZFN 1 and 2 (fourth bar). Based on these data a reliable comparison of the different TALEN pairs is not possible because the range of pH2A.X positive cells is close to the negative control.

The summarized data of three experiments in Figure 5.16 C document this observation. Untreated K562 cells showed about 2 % pH2A.X positive cells. Treatment with etoposide led to an increase to 57 % on average. Transfection with an empty expression vector (Strep-tag) delivered a signal of 13%, indicating a high background in this assay. ZFN1 and 2 showed a clearly increased signal of 33 % whereas the signals of all TALEN pairs (17, 13, 10, and 16 %) were in the range of the background signal (Strep-tag). The low sensitivity of the assay and the high background obscure the genotoxic potential of the different TALEN pairs and make it difficult to draw a reliable conclusion (Fig. 5.16 C).

### 5.6.2 The kinetic of the DNA damage response as a function of TALEN expression

TALEN-mediated genotoxicity in 5252 cells was assessed in a dose- and time-dependent manner. 5252 cells were induced by adding 5, 10, 50, and 100 ng/ml dox. After 24, 48, 72, and 96 hours the fraction of pH2A.X-positive cells, which is a surrogate marker of DNA double-strand breaks, was assessed by flow cytometry. The analysis was restricted to Cerulean-positive and therefore successfully induced cells. Since 5252 cells do not harbor the cognate sequence of TALEN 7 and TALEN 8 any DDR signal after dox-



**Fig. 5.17 Dox-mediated expression of TALEN 7 and 8 in 5252 cells induces a DNA damage signal**

**A.** Experimental overview: In 5252 cells expression of TALEN 7, TALEN 8 and Cerulean was induced by adding different concentrations of dox. After 24, 48, 72, and 96 hours post induction the cells were analyzed for the occurrence of the phosphorylated H2A.X (pH2A.X) histone, a marker for DNA double-strand breaks. In order to restrict the analysis to induced cells, only all cells analyzed were gated for the Cerulean-positive populations.

**B.** As an example, the fraction of pH2A.X positive 5252 cells induced with 100 ng/ml dox after 24, 48, 72, and 96 hours of one out of three experiments is shown here. 24 hours after induction of TALEN 7 and 8 expression the 5252 cells showed a clear signal of H2A.X phosphorylation (first panel, black line) which increased to 49 % 72 hours after induction (second and third panel, black lines). 96 hours after induction cells showed a reduced pH2A.X signal indicative of the reduction of TALEN expression (fourth panel, black line) as seen in Figure 5.7. In the absence of dox only a weak signal of pH2A.X was observed (grey area).

**C.** The results of three single, independent experiments are shown here. The first experiment shows a maximum of H2A.X phosphorylation 72 hours after addition of dox. A higher pH2A.X signal was observed with higher dox concentrations (left panel). The second experiment showed no clear time dependent

H2A.X phosphorylation. 5 ng/ml and 10 ng/ml of dox (dashed grey lines) delivered a higher signal than 50 ng/ml dox (solid black line with triangles) indicating that signal intensity did not correlate with different dox concentrations (middle panel). The highest signal in the third experiment was seen already after 24 hours of dox induction. The experiment showed a correlation of dox concentration and H2A.X phosphorylation (right panel) comparable to the first experiment (left panel).  
dox= doxycycline; pH2A.X= phosphorylated histone H2A.X

mediated induction of TALENs likely originates from unspecific cleavage events (Fig. 5.17 A). After adding 100 ng/ml dox to the 5252 cells the fraction of pH2A.X-positive cells was found to be 29, 39, 49, and 35 % 24, 48, 72, and 96 hours post induction, respectively. This finding indicates an obvious TALEN-mediated DNA damage response (Fig. 5.17 B). The three independent experiments yielded heterogeneous results, which made it impossible to define a time- or dose-dependent TALEN-mediated effect on the cellular DDR (Fig. 5. 17 C).

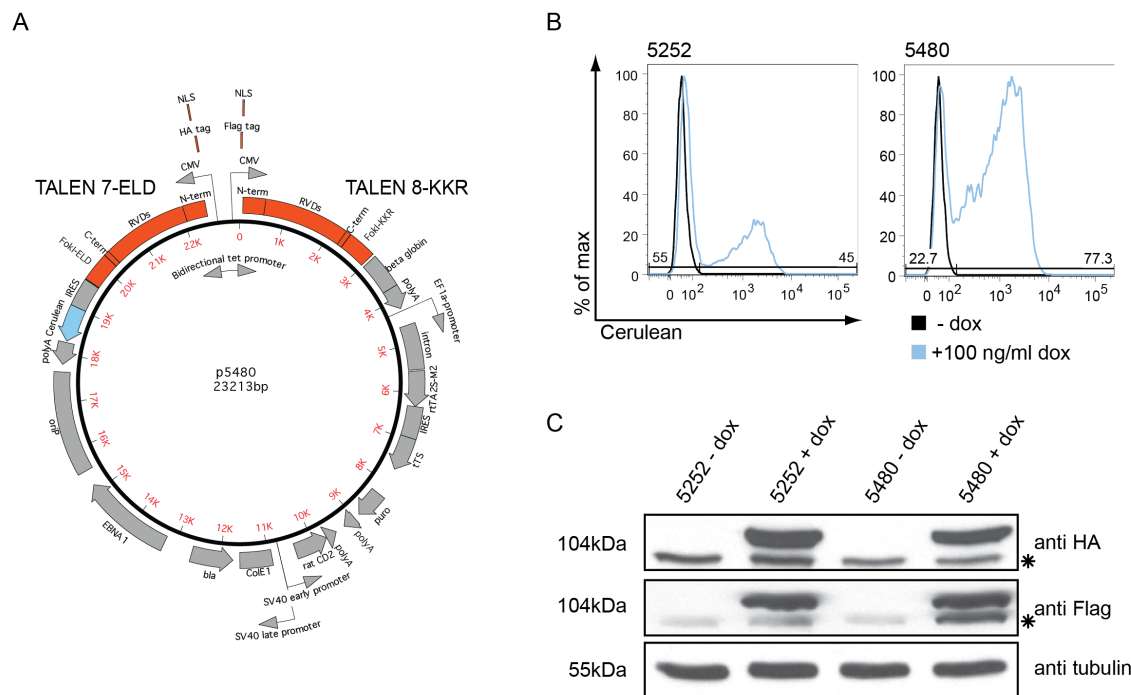
## 5.7 Application of TALENs with mutated FokI domains

After generating and testing several TALENs with different FokI mutants the TALEN 7-ELD and TALEN 8-KKR pair was used for further experimentation. No difference in the genotoxicity between the different TALENs was obvious in the experiment shown but this pair was chosen because it had the highest cleavage activity among TALENs with mutated FokI domains (Fig. 5.14, 5.15).

### 5.7.1 Establishment and analysis of a cell line expressing TALENs with mutated FokI domains

In order to test TALEN 7-ELD and TALEN 8-KKR in plaque assays an inducible cell line had to be generated. HEK293 cells were stably transfected with the oriP-based expression plasmid p5480 leading to the cell line 5480. 5480 is similar to 5252 except that the wild type FokI domains of TALEN 7 and TALEN 8 were replaced by the ELD and KKR mutants, respectively (Fig. 5.18 A). The newly established 5480 cell line and the previous 5252 cell line, which acts as a control were induced with 100 ng/ml dox and the fraction of Cerulean-positive cells was comparatively assessed. The newly generated

5480 cell line responded to dox addition; 77 % of the cells turned Cerulean-positive (Fig. 5.18 B). Western blot analysis revealed that TALEN 7 and TALEN 8 and TALEN 7-ELD and TALEN 8-KKR are expressed in 5252 cells and in the new cell line 5480 upon induction with dox (Fig. 5.18 C).



**Fig. 5.18 Generation of an inducible cell line for the expression of TALEN 7-ELD, TALEN 8-KKR and Cerulean**

HEK293 cells were stably transfected with an oriP-based expression plasmid. The plasmid constitutively expresses the surface marker CD2 and harbors a bidirectional doxycycline-inducible minimal CMV promoter conferring conditional expression of the TALENs and the phenotypical marker Cerulean.

**A.** The cell line 5480 is similar to the cell line 5252 (Fig. 5.3), except that the wild type FokI domains were replaced by the mutated versions ELD and KKR in TALEN 7 and TALEN 8, respectively. As indicated, the left ORF codes for TALEN 7-ELD fused to a NLS and a HA tag; the phenotypical marker Cerulean is expressed via an IRES element. The right ORF codes for TALEN 8-KKR fused to a NLS and a Flag tag. The HA and Flag tags confer independent detection of TALEN 7-ELD and 8-KKR expression levels by Western blot immunostaining.

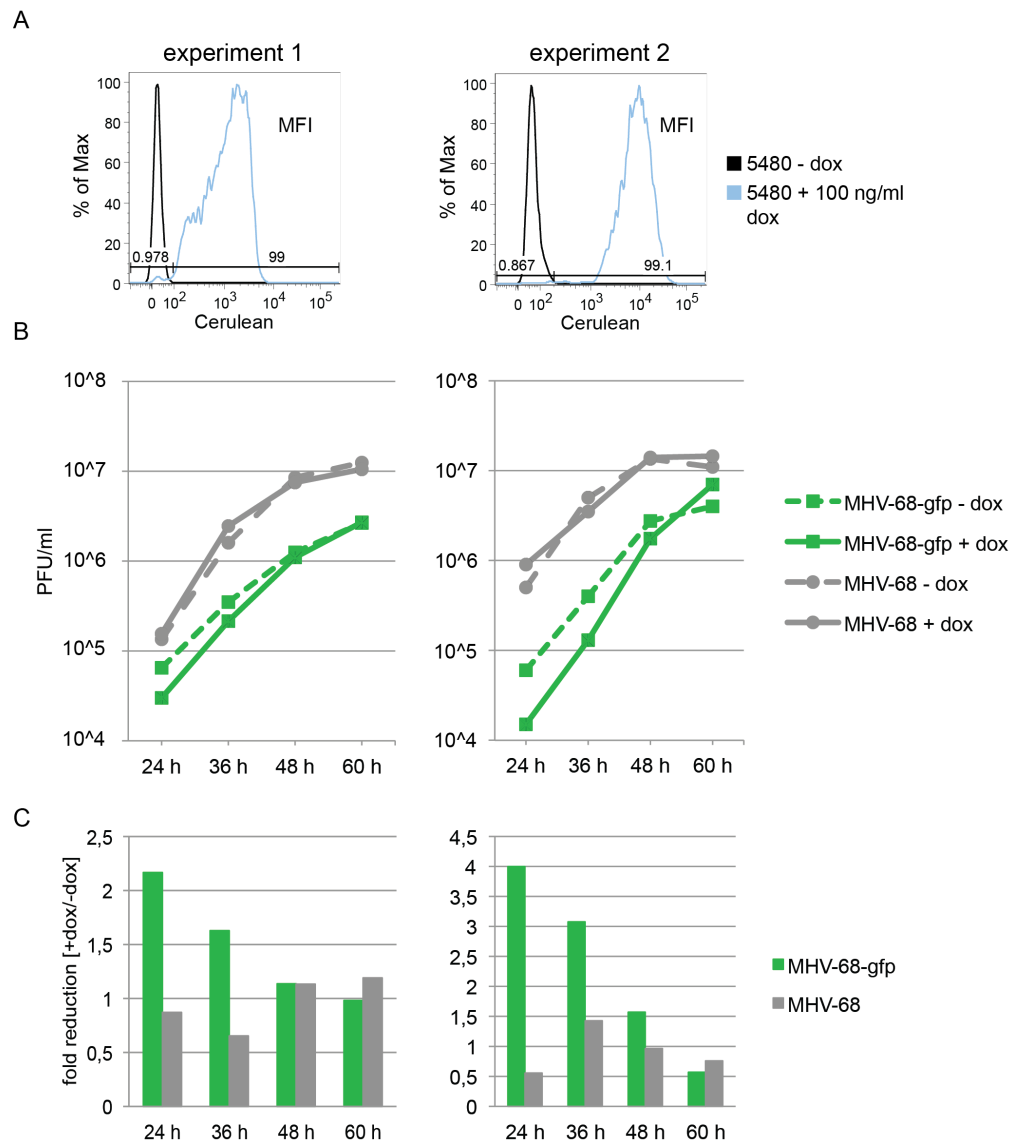
**B.** HEK293 cells stably transfected with the plasmid p5252 or p5480 were compared for Cerulean expression by flow cytometry with and without dox. The addition of dox led to 45 % or 77 % of Cerulean-positive cells (blue line) in 5252 or 5480 cells, respectively.

**C.** TALEN expression in the 5252 and 5480 cell lines without and with dox was analyzed by Western blot immunostaining. In the top panel an  $\alpha$ -HA antibody was used. The 5252 and 5480 cell lines without dox showed no signals but the addition of dox resulted in the appearance of a distinct band indicative of the expression of TALEN 7 or TALEN 7-ELD, respectively. The immunostaining with an  $\alpha$ -Flag antibody (middle panel) shows that the 5252 and 5480 cell lines delivered signals upon addition of dox indicative of the expressions of TALEN 8 and TALEN 8-KKR, respectively. The staining with an  $\alpha$ -tubulin antibody (bottom panel) demonstrates the equal loading of the different protein extracts. The asterisks point to unspecific signals.

dox= doxycycline, NLS= nuclear localization signal

### 5.7.2 MHV-68 and MHV-68-gfp *de novo* synthesis in cells expressing TALENs with mutated FokI domains

To assess the potential of TALEN 7-ELD and TALEN 8-KKR to prevent viral synthesis two plaque assays were performed. Similarly as described in Figure 5.12, 5480 cells were induced with 100 ng/ml dox and twelve hours later sorted for Cerulean expression. Prior to infection the induction rate was assessed by flow cytometry showing induction rates of 99 % in two independent experiments. The second experiment revealed a higher mean fluorescence intensity of Cerulean indicating a higher TALEN expression level (Fig. 5.19 A). The induced and sorted cells were infected with MHV-68-gfp and MHV-68. The latter lacks the *gfp* coding sequence and therefore the cognate sequence of TALEN 7-ELD and TALEN 8-KKR. Therefore, MHV-68 synthesis should not be affected by the TALENs expressed in 5480 cells. In contrast, MHV-68-gfp harbors the target sequence of TALEN 7-ELD and TALEN 8-KKR and therefore this virus should be under control of the TALENs. In the two experiments the control virus MHV-68 replicated similarly in uninduced and induced cells (Fig. 5.19 B) indicating a reduced cytotoxic effect of TALENs with mutated FokI domains compared to TALENs with wild type FokI domains in 5252 cells (Fig. 5.12). Plaque assays with MHV-68-gfp revealed no or only a weak reduction of viral *de novo* synthesis (Fig. 5.19 B). The slight effect in the second experiment could be due to the better induction compared to the first experiment (Fig. 5.19 A). The calculation of the ratios +/- dox revealed no influence of TALEN expression on *de novo* synthesis of MHV-68 virus indicated by a ratio at about one at all time points analyzed. Infection with MHV-68-gfp showed a weak effect after 24 hours in the second experiment, only; no effect was observed in the first experiment. These findings suggested that TALEN 7-ELD and TALEN 8-KKR might be more specific than wild type TALEN 7 and TALEN 8 but their reduced cleavage activity does not limit viral *de novo* synthesis of MHV-68-gfp efficiently (Fig. 5.19 C).



**Fig. 5.19 The combination of TALEN 7-ELD and TALEN 8-KKR is not able to repress MHV-68-gfp *de novo* synthesis**

Cells stably transfected with the oriP-derived expression plasmid p5480, which expresses TALEN 7-ELD and 8-KKR upon addition of dox were infected with two different viruses. MHV-68-gfp contains the *gfp* sequence, which is the target of TALEN 7-ELD and 8-KKR. The second virus MHV-68 lacks the TALEN 7-ELD and 8-KKR target sequence. The concentration of both viruses in 5480 cell supernatants without or with dox was compared in plaque assays. The experiment was performed as described in Figure 5.12; two independent experiments are shown here.

**A.** The fraction of Cerulean-positive cells was assessed after addition of dox and cell sorting. In both experiments 99 % of the cells were induced but in the second experiment the mean fluorescent intensity of the Cerulean-positive population was higher (right panel) indicating a higher expression of TALEN 7-ELD and TALEN 8-KKR.

**B.** In the two independent experiments *de novo* synthesis of the MHV-68 virus was similar in the absence and presence of dox (dashed and solid grey lines). As compared to TALEN 7 and 8 shown in Fig. 5.12 C, TALEN 7-ELD and TALEN 8-KKR showed a reduced cytotoxicity. When induced 5480 cells (solid green line) were infected with MHV-68-gfp no reduction (left panel) or only a weak reduction (right panel, 24 hours) was observed compared to uninduced 5480 cells (dashed green line). This result is in line with a reduced cleavage activity of TALEN 7-ELD and TALEN 8-KKR compared to TALEN 7 and 8 (Fig. 5.14 and Fig. 5.15), which does not impair MHV-68-gfp *de novo* synthesis. The slight effect in the second experiment (right panel) might be due to a higher expression of TALEN 7-ELD and 8-KKR as compared to the first experiment (left panels).

**C.** The ratios of virus concentrations obtained from induced and uninduced 5480 cell supernatants were calculated. TALEN 7-ELD and TALEN 8-KKR (grey bars) did not impair synthesis of MHV-68 in both



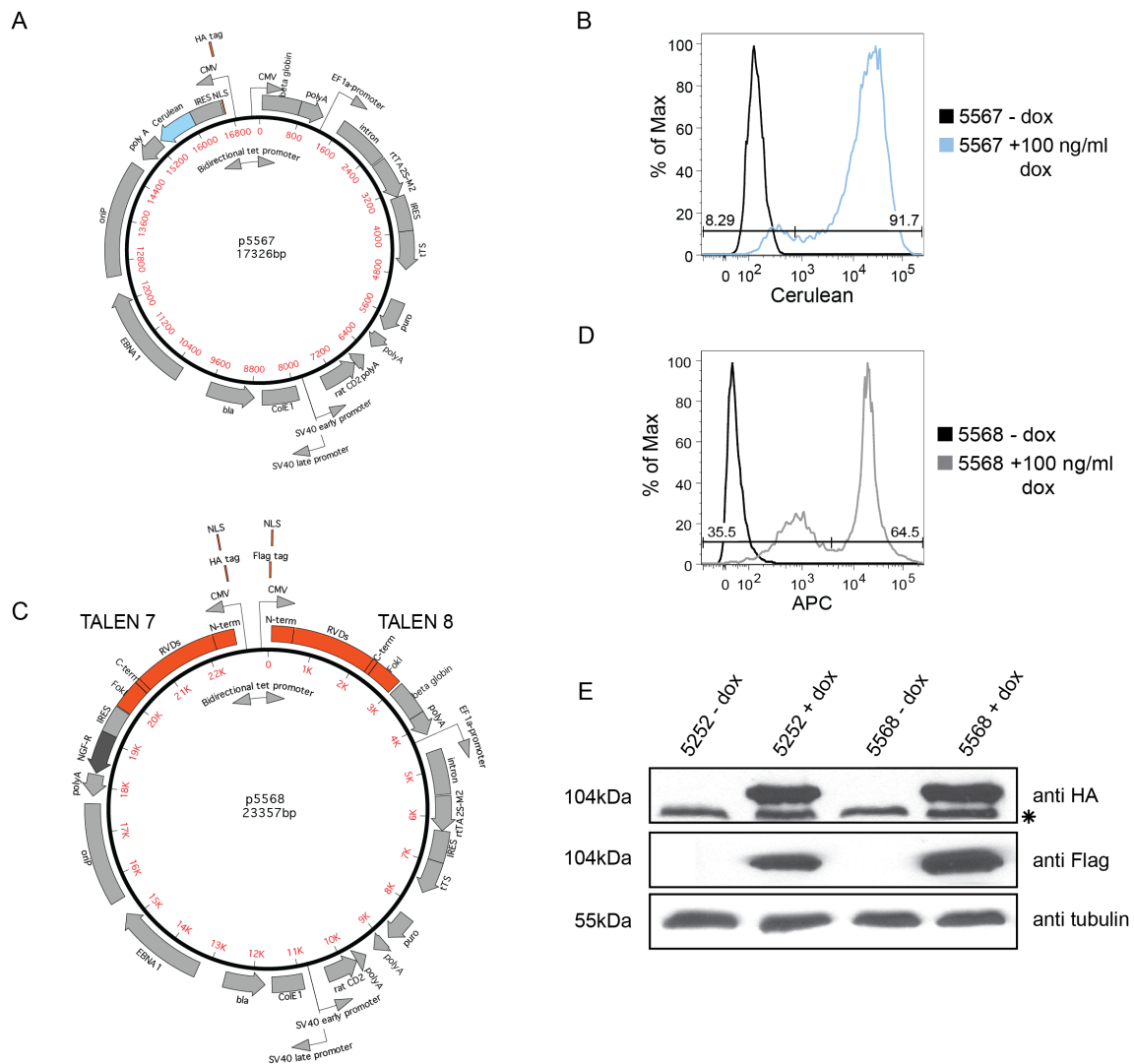
experiments indicated by a ratio of one. In the first experiment (left panel) no clear reduction of MHV-68-gfp was achieved. In the second experiment (right panel) MHV-68-gfp *de novo* synthesis was slightly reduced 24 hours after infection, only, suggesting that the specificity of cleavage of TALEN 7-ELD and TALEN 8-KKR is improved at the expense of their cleavage activity.  
dox= doxycycline, PFU= plaque forming units

## 5.8 Effect of induced TALEN expression on cellular proliferation

Several experiments have shown or suggested that TALENs have cytotoxic side effects. Figure 5.7 revealed that the percentage of TALEN expressing 5252 cells continuously decreased after induction whereas mCherry expression increased to reach almost 100 % of induced cells. 5252 cells infected with MHV-68 showed an unspecific reduction of viral *de novo* synthesis (Fig. 5.13) but this effect was not observed in 5480 cells, which express the more specific mutant TALEN 7-ELD and TALEN 8-KKR (Fig. 5.19). Experiments with K562 cells stained for the phosphorylated version of histone H2A.X, a marker of DNA damage, were not conclusive due to the high background of the assay (Fig. 5.16). Staining of 5252 cells for pH2A.X after induced TALEN expression clearly showed a TALEN-mediated genotoxic effect (Fig. 5.17). To further address this issue the effect of TALEN expression on cellular proliferation was assessed.

### 5.8.1 Establishment of two additional inducible cell lines

Two additional inducible cell lines were generated, which served as controls in the following proliferation experiment. The oriP-based expression plasmid p5567 is similar to p5252, except that the coding sequences for TALEN 7 and 8 were removed. Therefore, 5567 cells express Cerulean upon induction, only. This cell line was used as an additional negative control in the proliferation assay to exclude effects of Cerulean expression on cellular proliferation (Fig. 5.20 A). 91 % of HEK293 cells stably transfected with p5567 turned Cerulean-positive after the addition of dox (Fig. 5.20 B).



**Fig. 5.20 Generation of two cell lines for the conditional expression of TALEN 7, TALEN 8 and NGF-R or Cerulean**

HEK293 cells were stably transfected with two oriP-based expression plasmids. Both plasmids constitutively express the surface marker CD2 and harbor a bidirectional doxycycline-inducible minimal CMV promoter conferring conditional expression.

**A.** HEK293 cells stably transfected with the plasmid p5567 led to the cell line 5567, which expresses Cerulean upon dox-mediated induction. This cell line is similar to the cell line 5252 (Fig. 5.3), which however, expresses the coding sequences for TALEN 7 and TALEN 8 in conjunction with Cerulean upon induction.

**B.** Cerulean expression in HEK293 cells stably transfected with the plasmid p5567 were compared by flow cytometry. The addition of dox led to 91 % of Cerulean-positive cells (blue line); Cerulean was not detectably expressed in the absence of dox (black line).

**C.** The cell line 5568 was generated by stably transfecting HEK293 cells with the plasmid p5568. It is similar to p5252 (Fig. 5.3) but the phenotypical marker Cerulean in p5252 was exchanged for the surface molecule NGF-R in p5568. Addition of dox leads to the expression of TALEN 7 fused to a NLS and a HA tag, the surface marker NGF-R linked via an IRES element (left ORF) and TALEN 8 fused to a NLS and a Flag tag (right ORF). Fusion of the different HA and Flag tags allows the independent detection in Western blot immunostaining of TALEN 7 and TALEN 8, respectively.

**D.** HEK293 cells stably transfected with p5568 were analyzed by flow cytometry to evaluate the expression of NGF-R. After the addition of dox 64 % of the cells turned positive expressing NGF-R (grey line).

**E.** The 5252 and 5568 cell lines without and with dox were analyzed by Western blot immunostaining analyzing the expression of TALEN 7 and 8. In the top panel an  $\alpha$ -HA antibody was used. The 5252 and 5568 cell lines without dox showed no TALEN-specific signals but the addition of dox resulted in the appearance of a distinct band indicative of the expression of TALEN 7 in both cell lines. Immunostaining

with an  $\alpha$ -Flag antibody (middle panel) shows that the 5252 and 5568 cell lines showed signals upon addition of dox indicative of the expression of TALEN 8. The staining with an  $\alpha$ -tubulin antibody (bottom panel) demonstrates the equal loading of the different protein extracts. The asterisk points to a background signal unrelated to the HA-tag.

dox= doxycycline, NLS= nuclear localization signal

HEK293 cells stably transfected with p5568 express TALEN 7 and TALEN 8 similar to 5252 cells but the surrogate marker Cerulean was replaced by the surface molecule NGF-R. The 5568 cell line served as a control to exclude any effects that might arise from an interference of the TALENs with Cerulean or its coding sequence, which is related to *gfp* (Fig. 5.20 C). After induction with dox 64 % of the 5568 cells turned NGF-R-positive (Fig. 5.20 D). Western blot analysis revealed an expression of TALEN 7 and TALEN 8 in 5568 cells similar to 5252 cells (Fig. 5.20 E).

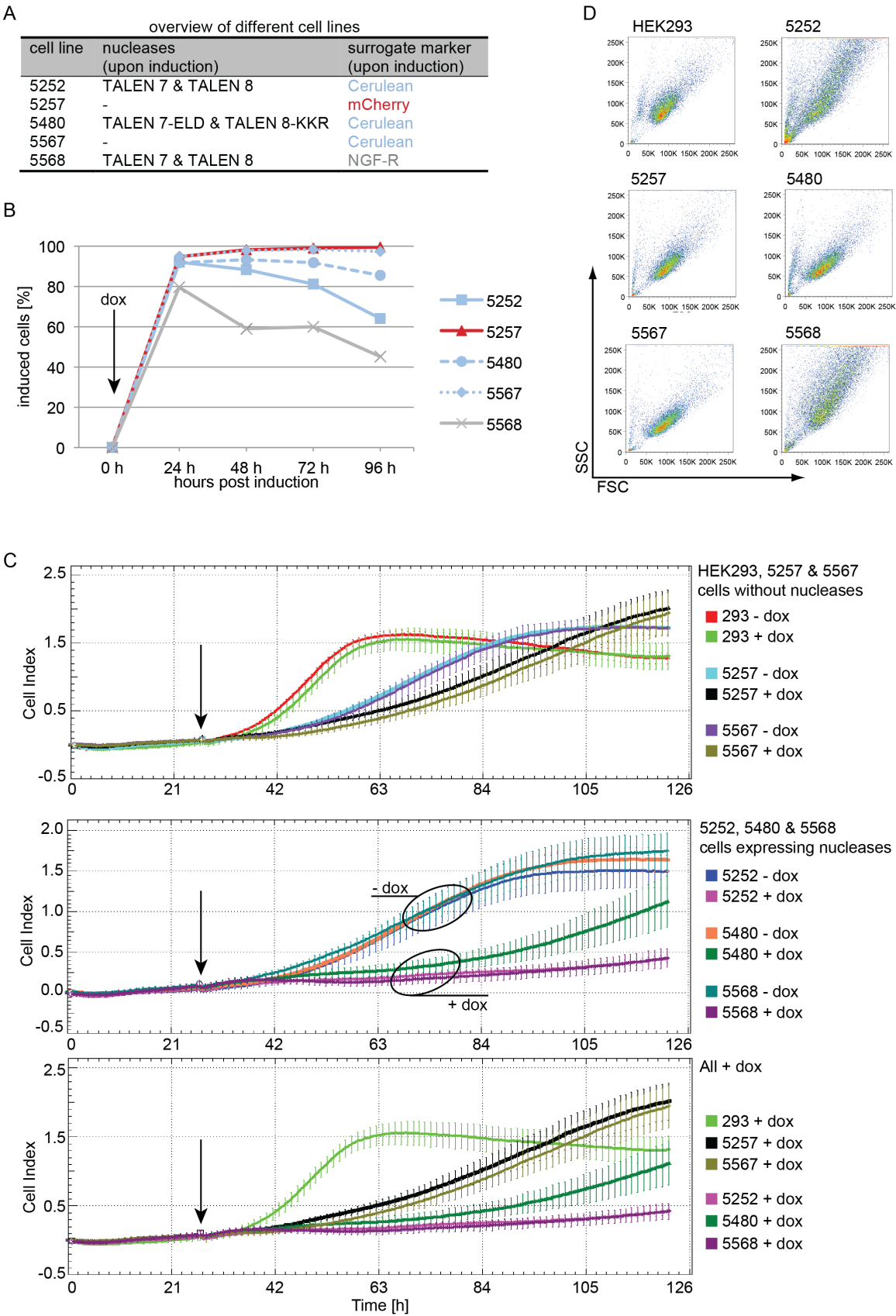
### 5.8.2 Proliferation analysis of inducible cell lines

Proliferations of five different cell lines were measured and compared using the Xcelligence instrument. Both, 5252 and 5568 express TALEN 7 and TALEN 8 but differ in the phenotypical markers Cerulean (5252) and NGF-R (5568). The cell line 5480 expresses TALEN 7-ELD and TALEN 8-KKR. The two cell lines 5257 and 5567 do not express any nuclease but the phenotypical markers mCherry and Cerulean, respectively (Fig. 5.21 A). 24 hours after seeding the cells protein expression was induced by adding dox (black arrow). The induced expression of the phenotypic markers in all cell lines was monitored during the course of the experiment, because ectopic expression is a crucial parameter and a measure of cells expressing the different nucleases. The cell lines 5257 and 5567 express the phenotypical markers, only. They showed a high induction rate of more than 90 % 24 hours after induction further increasing to almost 100 % in the course of the experiment. The two TALEN 7 and 8 expressing cell lines 5252 and 5568 showed a maximum induction rate 24 hours after induction, which declined over time (Fig. 5.21 B). The 5480 cell line expresses TALEN 7-ELD and TALEN 8-KKR with the mutated FokI domains. It showed an intermediate phenotype with a stable rate of induction between 24 and 72 hours and a slight reduction at 96 hours. Together, these results indicated that HEK293 cells do not tolerate the constitutive

expression of wild type TALENs whereas the expression of TALEN 7-ELD and TALEN 8-KKR seems to be less cytotoxic (Fig. 5.21 B).

With the aid of the Xcelligence instrument I monitored the rate of cellular proliferation over time. HEK293 cells were not affected by 100 ng/ml dox and proliferated undisturbed by the compound (Fig. 5.21 C). The 5257 and 5567 control cells showed a slight reduction of proliferation upon the addition of dox (black arrow), and all TALEN cell lines showed the same proliferation in the absence of dox. Upon induction, the wild type TALEN expressing cell lines (5252 and 5568) showed a dramatic decrease of proliferation (Fig. 5.21 C). Comparing the two controls (5257 and 5567) and the TALEN expressing cells (5252 and 5568) my observations so far suggested that the induced expression of the TALENs and not the marker proteins caused the reduced rate of cellular proliferation. The expression of TALEN 7-ELD and TALEN 8-KKR revealed an intermediate phenotype as expected (Fig. 5.21 C), because mutations in the FokI domains were shown to reduce the genotoxic effects. A comparison of all cell lines in the presence of dox shows a strong negative effect of TALEN expression on the rate of cellular proliferation whereas in cells expressing mCherry (5257) or Cerulean (5567), only, this effect is less obvious. The parental HEK293 cells are not influenced by the addition of dox (Fig. 5.21 C).

I also looked at the overall phenotypes of the cells by forward and sideward scatter criteria on a FACS instrument. The cell lines, which did not express any nucleases showed distinct populations indicative of a regular phenotype (HEK293, 5257, and 5567 cells in Fig. 5.21 D). In contrast, the wild type TALEN expressing cell lines (5252 and 5568) showed a very unusual, heterogeneous population in the FSC/SSC plot indicating that TALEN expression had an impact on cell size and granularity. This effect was less pronounced upon induction of mutated TALEN 7-ELD and TALEN 8-KKR suggesting reduced cytotoxic effects (5480 cells in Fig. 5.21 D).



**Fig. 5.21 Expression of TALEN 7 and 8 interferes with cellular proliferation**  
Proliferation of HEK293 cells stably transfected with different oriP-based expression vectors was compared without and with dox-mediated induction of protein expression using the Xcelligence device. Induction rates were determined by flow cytometry at the indicated intervals.

**A.** To ensure similar conditions, HEK293 cells were transfected with the five different plasmids, selected and sorted for CD2 expression, which correlates with high rates of induced protein expression (Fig. 5.6). The table provides an overview of the five cell lines. Two cell lines were negative controls and expressed mCherry (5257) or Cerulean (5567), upon induction, only. Two cell lines expressed TALEN 7 and 8 upon dox-mediated induction but co-express different phenotypical markers Cerulean (5252) or NGF-R (5568). In addition, the effect of mutated FokI domains in TALEN 7-ELD and TALEN 8-KKR was assessed in a fifth cell line (5480).

**B.** 24 hours after seeding the cells dox was added (black arrow) and the proportion of induced cells was assessed by flow cytometry to verify the fraction of cells expressing the induced proteins. The cell lines 5257 (mCherry only – solid red line) and 5567 (Cerulean only – dotted blue line) showed an induction rate of more than 90 % 24 hours after addition of dox, which increased even further in the course of the experiment. The TALEN 7 and 8 expressing cell lines with the phenotypical markers Cerulean (solid blue line) or NGF-R (solid grey line) reached a maximum at 24 hours after adding dox, which declined over time. Expression of TALEN 7-ELD and 8-KKR showed an intermediate phenotype with a stable fraction of induced cells from 24 to 72 hours and a slight reduction at 96 hours (dashed blue line).

**C.** Cellular proliferation was assessed by a Xcelligence device, which measures the impedance of the cells as a function of proliferation and spreading of adherent cells on electrodes situated on the bottom of the wells. The changes in impedance are expressed as cell index. Proliferation of cell lines in the presence of dox (black arrow) and the absence of dox was compared. In the upper panel control cells without nucleases are shown. Parental HEK293 cells showed the same proliferation characteristic in the absence and presence of dox (red and green lines) as expected. Uninduced 5257 and 5567 cells proliferated slower compared to parental HEK293 cells (turquoise and violet lines). The addition of dox led to the expression of mCherry (5257) or Cerulean (5567) and a slightly reduced proliferation (black and dark yellow lines). In the middle panel cell lines, which express different TALENs upon induction are compared. When dox was omitted the three cell lines showed similar proliferation rates. When expression of TALEN 7 and 8 was induced (5252 and 5568) the proliferation of the cells was dramatically reduced (pink and dark violet lines) indicating that cellular proliferation and the expression of TALEN 7 and 8 are incompatible. This phenotype was observed in cells co-expressing the combination of TALEN 7, TALEN 8 and Cerulean (5252, pink) or cells co-expressing TALEN 7, TALEN 8 and NGF-R (5568, dark violet). From these observations it can be concluded that the expression of the phenotypical markers Cerulean (5252) or NGF-R (5568) do not interfere with proliferation but TALEN 7 and 8. When TALEN 7-ELD and TALEN 8-KKR were induced (5480) an intermediate phenotype was observed (dark green line) indicating that the mutated FokI domains led to reduced cytotoxic effects as expected. A comparison of all the samples with dox (lower panel) shows that cells that exclusively express mCherry (5257) or Cerulean (5567) (black and dark yellow lines) show a slightly impaired proliferation rate compared to parental HEK293 (green line) cells. Reduced proliferation was most obvious in cells expressing the TALENs (dark green, pink and dark violet lines).

**D.** 96 hours after dox induction cells were analyzed by flow cytometry. Here the FSC / SSC dot plots of the six different cell lines are shown. In the left column the cell lines are depicted, which express no additional protein upon induction (HEK293) or express mCherry (5257) or Cerulean (5567), only. In these samples flow cytometry showed distinct populations whereas an expression of TALENs results in heterogeneous population (right panels). This effect was more pronounced in cells expressing TALEN 7 and 8 (5252 and 5568) than in TALEN 7-ELD and 8-KKR expressing cells (5480). This observation indicates that the induced expression of TALENs influences cell size and cellular granularity.

dox= doxycycline

## 6. Discussion

In 2009 two groups deciphered the code of DNA recognition by TAL effectors (BOCH *et al.* 2009; MOSCOU and BOGDANOVE 2009). Since then, TAL effectors have emerged as versatile tools for the easy and rapid generation of DNA-binding proteins with novel user-defined sequence specificities. By replacing the transcriptional activation domain of the TALE with the cleavage domain of the FokI restriction enzyme, TALE nucleases (TALENs) have been constructed (CHRISTIAN *et al.* 2010; LI *et al.* 2011; MAHFOUZ *et al.* 2011). The modularity of the DNA binding domain facilitates the rapid construction of nucleases with new target specificities making the application of TALENs accessible for many researchers. With the aid of new cloning methods such as ligation-independent cloning the potential of generating TALEN with novel target specificities has become even less cumbersome or time consuming (SCHMID-BURGK *et al.* 2013). TALENs were used in the field of genomic engineering by exploiting cellular DNA repair mechanisms: non-homologous end joining and homologous recombination. Using these techniques, endogenous loci have been edited in plants (CHRISTIAN *et al.* 2010), yeast (LI *et al.* 2011), zebrafish (DAHLEM *et al.* 2012; ZU *et al.* 2013), drosophila (LIU *et al.* 2012), rat (TESSON *et al.* 2011), and human somatic (MILLER *et al.* 2007) and stem cells (HOCKEMEYER *et al.* 2011). In the framework of this thesis this new technology has been adopted and a new application was developed.

### 6.1 Scope of this project

The scope of this thesis was to develop a new application of TALENs with inherent antiviral functions. In this system TALENs would target and cleave a sequence present in the genome of double-stranded DNA viruses but absent in the genome of the viral host cell. The working hypothesis of the thesis built on the assumption that a TALEN-

mediated DNA DSB within the viral genome will inhibit or abrogate viral replication. Cells expressing TALENs, which target a sequence contained in the viral genome would therefore acquire a state of endogenous resistance against the virus. This simple strategy is appealing since it builds on the genomic viral sequence, only, but does not require a fundamental understanding of the virus lifestyle. The murine herpes virus MHV-68-gfp harboring the *gfp* coding sequence was used as a model system.

The initial step of the project was to define the DNA sequence of interest in the viral genome. We decided to use the *gfp* gene for different reasons. First, TALENs targeting the *gfp* coding sequence could be used with any DNA virus harboring the *gfp* gene. For instance, infection experiments with the Epstein-Barr virus would be technically feasible (DELECLUSE *et al.* 1998). Additionally, the function of GFP as a phenotypical marker provides an easy read-out in *gfp* knockout experiments. We assumed that cells expressing the TALENs would not turn GFP-positive as the *gfp* sequence is cleaved early after infection. This aspect was particularly important in the beginning of the project as the loss of *gfp* provides a simple readout in infection experiments. Targeting the *gfp* sequence also facilitated comparing viral *de novo* synthesis of two nearly identical viruses MHV-68 and MHV-68-gfp, in TALEN expressing cells. Since these viruses differ in the presence or absence of the *gfp* and therefore the TALEN target sequence, only, the comparison of *de novo* synthesis of both viruses allows discriminating between specific and unspecific TALEN-mediated effects.

A possible detrimental aspect of the strategy is the fact that a TALEN-mediated DSB within the *gfp* coding sequence can be repaired by the error prone NHEJ leading to a functional knockout of the gene. Since *gfp* is not essential, it is conceivable that its mutation will not prevent viral replication. Upon NHEJ the TALENs cognate sequence within the viral genome might become modified or is lost, resulting in a virus resistant to TALENs. This evasive scenario could be circumvented by targeting one of MHV-68's 41 genes that are essential for *de novo* synthesis *in vitro* (SONG *et al.* 2005). Here NHEJ repair after a TALEN-mediated DNA DSB would most likely result in a lethal mutation of the essential gene inhibiting or preventing viral replication. The viral transcriptional activator Rta would be an interesting target in this context because a knockout of this gene will inhibit the lytic replication of MHV-68 (WU *et al.* 2001).



The first aim of the thesis was the construction of two TALEN pairs targeting a sequence within the GFP coding sequence and to compare their cleavage activity *in vitro* and *in vivo*. After choosing the TALEN pair with the highest cleavage activity I addressed the question whether TALENs cleave sequences within the viral genome at the level of DNA or protein expression. Whether a double-strand break within the viral genome would impair viral replication was the central question of the thesis. Therefore, I assessed and compared the viral *de novo* synthesis in control cells or cell lines expressing the TALENs. Constitutive expression of the TALENs, presumably conferring an intrinsic viral resistance to the cells, could have undesirable side effects such as genotoxicity. Since this approach has not been tried before I addressed the issue of possible TALEN-induced side-effects in several experiments.

## 6.2 Biochemical analysis of ZFN and TALE DNA binding

### 6.2.1 ZFN

Custom-nucleases, in particular ZFNs, are widely used in the field of genomic engineering. Yet, not much is known about their DNA binding and dimerization characteristics. Consequently, in the framework of this thesis the biochemical characterization of DNA binding and dimerization of ZFNs was analyzed. Additionally, I compared the DNA binding characteristics of a synthetic TALE with a natural TALE of *Xanthomonas*.

In technical terms, I used electrophoretic mobility shift assays with ZFN monomers and dimers in my experiments. Dimerization of the FokI domain would result in the cleavage of the DNA template (BITINAITE *et al.* 1998), which would be detrimental for the experimental outcome. Therefore, I constructed ZFNs with inactivated FokI nuclease domains by introducing a single amino acid substitution (D450A) in the catalytic center of the nuclease domain. I denoted the ZFNs with inactivated nuclease domains ZFΔN. EMSA experiments of single ZFΔNs revealed  $K_d$  values of  $279 \pm 36$  nM and  $293 \pm 89$  nM,

respectively. Several publications cover this aspect and reveal  $K_d$  values for zinc finger proteins ranging from the low  $\mu\text{M}$  to the low  $\text{pM}$  range (DESJARLAIS and BERG 1993; LIU *et al.* 1997; KIM and PABO 1998; HURT *et al.* 2003; SHIMIZU *et al.* 2011; SCHIERLING *et al.* 2012; BHAKTA *et al.* 2013). The extremely wide range of zinc finger DNA binding affinities is puzzling. This heterogeneous picture might result from different experimental readouts or suggests that the design of zinc fingers has an major impact not only on the specificity but also the affinity of DNA binding.

Next, I wanted to assess the  $K_d$  value for a dimer consisting of ZF $\Delta$ N 1 and ZF $\Delta$ N 2. Unexpectedly, only a small fraction of DNA fragments was bound by a ZF $\Delta$ N dimer making the calculation of the  $K_d$  value impossible (Fig. 4.4). This observation indicates a low stability of the dimer itself and suggests that a ZF $\Delta$ N pair on DNA is very unstable and occurs transiently, only. This is in line with published experiments, which document a weak FokI interaction resulting in FokI monomers in solution even at micro-molar concentrations (VANAMEE *et al.* 2001; CATTO *et al.* 2006). Additionally, dimerization characteristic of ZF $\Delta$ Ns with mutated FokI domains, ZF $\Delta$ N 1-RR and ZF $\Delta$ N 2-DD, was analyzed. These mutated FokI domains allow heterodimer formation, only, and therefore increase cleavage specificity of the ZFN pair. EMSAs revealed that dimerization of the ZF $\Delta$ Ns with mutated FokI domains was even less obvious compared to ZF $\Delta$ Ns with wild type FokI domains. This outcome was expected as the mutations in the FokI domains are known to lower the dimerization energy (SZCZEPEK *et al.* 2007). Again, a  $K_d$  value for the ZF $\Delta$ N dimer with mutated FokI domains could not be calculated (Fig. 4.4). EMSA experiments with ZF $\Delta$ N 2 revealed that the template size has no influence on DNA binding affinity, but several bands appeared with increasing protein concentration in EMSAs with a longer template. Multiple bands indicate unspecific binding of ZF $\Delta$ N 2 to DNA. This observation was made at protein concentrations much higher than in an *in vivo* situation suggesting that unspecific binding of ZFN 2 *in vivo* might be less frequent (Fig. 4.5).

### 6.2.2 TALE

The DNA binding characteristics of the synthetic TALE 5 was assessed in EMSA experiments using two DNA templates of different size. The binding affinity of TALE 5 to the shorter DNA template was too weak to calculate the corresponding  $K_d$  value whereas the binding affinity was much higher with a DNA template of 706 bp in length resulting in a  $K_d$  value of  $301 \pm 21$  nM. EMSA experiments with the natural TALE HAX3 and its cognate DNA binding motif showed higher binding affinities compared to the synthetic TALE 5 indicating that the design of constructed TALEs can affect binding affinity.  $K_d$  values of  $674 \pm 139$  nM and  $76 \pm 30$  nM were calculated for the shorter and the longer template, respectively, confirming the observation that TALEs bind longer templates with considerably higher affinity (Fig. 4.6). It has been shown that the *lac* repressor slides on chromosomal DNA in order to rapidly find and bind to its specific sequence (HAMMAR *et al.* 2012). It is conceivable that TALEs also utilize such a scanning process to identify their target sites rapidly and accurately. A scanning mode would explain the different binding affinities of templates differing in size: a TALE molecule could easily drop off a shorter template during sliding.

Several groups have addressed the question of DNA binding characteristics.  $K_d$  values ranging from the low  $\mu$ M to the high pM range were suggested (BULTMANN *et al.* 2012; CHRISTIAN *et al.* 2012; GAO *et al.* 2012; MECKLER *et al.* 2013). Several different parameters influence the DNA binding affinity of a TALE. It became apparent that more than 94 residues of the TALE N-terminus upstream of the repeat domain are required for high-affinity binding to DNA whereas the C-terminus seems to be dispensable for DNA binding (MECKLER *et al.* 2013). The repeat domain without the N-terminus shows no DNA binding at all whereas the N-terminus itself binds unspecifically to DNA with a very low affinity (GAO *et al.* 2012). The affinity of the TALE protein towards its cognate sequence strongly depends on the RVD composition of the repeat domain. The relative affinities were ranked: NG > HD/NN >> NI > NK. The NN RVD binds both A and G but the latter with higher affinity. Additionally, the RVD composition as well as the sequence context affect the affinity of a TALE protein to its cognate sequence (CHRISTIAN *et al.* 2012; MECKLER *et al.* 2013). In order to validate our results we assessed the  $K_d$  value of the prototypical AP-1 heterodimer c-Jun/c-Fos and achieved results comparable to literature (Fig. 4.7; JOHN *et al.* 1996; CHYTIL *et al.* 1998).

## 6.3 TALENs as antiviral factors

### 6.3.1 Generation of *gfp*-specific TALENs

The first aim of the project was the generation of *gfp*-specific TALEN pairs. In a collaboration with Thomas Lahaye and Orlando de Lange, we constructed two TALEN pairs by modular cloning (Fig. 4.1; MORBITZER *et al.* 2011). The nucleotide G can be targeted with two different repeats in the TALE DNA binding domain. We used NN RVDs to target this nucleotide because of the higher activity compared to NK repeats (CHRISTIAN *et al.* 2012). I utilized a TALEN backbone with truncated N-terminus and C-terminus, which was shown to have high cleavage activity but low cytotoxicity (MUSSOLINO *et al.* 2011). Mussolino *et al.* suggested an optimal spacer length for TALENs based on this scaffold of 12 or 13 bp. For this reason we designed TALEN 5 and 6 with a 12 bp and TALEN 7 and 8 with a 13 bp spacer (Fig. 4.2). Cleavage activities of both TALEN pairs were first assessed in an *in vitro* cleavage assay. Both TALEN pairs were expressed at the same level but TALEN 7 and 8 showed a higher cleavage activity (Fig. 5.1). Additionally, I assessed the cleavage activities in an *in vivo* assay. The target sequence was a single copy of the *gfp* coding sequence integrated in the cellular genome. In this assay the target sequence is localized in the nucleus, as would be the genome of a herpes virus upon cellular infection. In accordance with the *in vitro* assay TALEN 7 and 8 again performed better than TALEN 5 and 6 (Fig. 5.2).

At this point it was not possible to predict the required amount of cleavage events in a single cell to confer resistance to a virus. Two possible parameters have to be considered. First, the number of viral genomes in a single cell after infection is a critical point. Second, an already cleaved viral genome might be rejoined by the cellular repair machinery requiring yet another TALEN-mediated cleavage. Based on these results and assumptions we decided to continue the project with the TALEN 7 and 8 pair due to its superior cleavage characteristics.

### 6.3.2 Cleavage of viral genomes

With a highly active TALEN pair I continued with cleavage experiments monitoring the events after viral infection with MHV-68-gfp. This general approach required a cell line expressing TALEN 7 and 8, which is permissive for infection with MHV-68. We assumed that cells would not tolerate a constitutively high expression of the nucleases for a period of weeks or months because TALENs have been reported to mediate genotoxic effects (MUSSOLINO *et al.* 2011). We decided to establish a dox-inducible cell line to express TALENs at will (BORNKAMM *et al.* 2005). We chose HEK293 cells for the following experiments since they were shown to work with our oriP-based expression system and support lytic replication of the MHV-68 virus (unpublished data). We established two conditional cell lines, 5252 expressing TALEN 7, TALEN 8 and Cerulean and 5257, which served as a control cell line, and expresses the phenotypical marker mCherry upon induction.

An initial series of experiments revealed that in the cell line 5252, which expresses TALEN 7 and TALEN 8 only about 50 % of the cells responded to dox-mediated induction (Fig. 5.3). At this point I was facing two problems. (i) An induction rate of 50 % was too low for certain planned experiments and (ii) induction rates constantly declined when the cells were kept in culture without induction (Fig. 5.5 A). Therefore, I directed my efforts towards the optimization of the inducible system. As a first experiment had revealed dox concentration and induction rate in the bulk population did not correlate (Fig. 5.4) I decided to establish clonal cell lines. The clonal cell lines 5252.37 and 5257.40 showed higher induction rates with 75 % and 93 %, respectively, as compared to bulk transfected HEK293 cells. Yet, the induction rate of 75 % in 5252 cells was not high enough and again the induction rates in both cell lines were decreasing over time.

To understand the failure to induce gene expression in all cells, I investigated whether the unresponsive cells had lost the plasmid or their dox-mediated inducibility. 5252 cells transfected with the oriP-derived expression plasmid p5252 also express the surface marker CD2, which is under the control of a constitutive promoter on the same plasmid. Consequently, the surface marker CD2 could be used as an indicator of the presence of p5252 in the cell population. To address this issue 5252 cells were induced

and stained for the surface marker CD2. Surprisingly, only 31 % of the cells were CD2-positive although they all had been kept under selection. 17 % of the cells were CD2-negative but turned Cerulean-positive upon induction and 52 % of the cells were double-negative for both markers. These findings led to important conclusions: (i) The population of CD2-negative but Cerulean-positive population cannot be easily explained. All dox responsive cells harboring the plasmid should be CD2-positive but there are obviously cells stably transfected with p5252, which are CD2 negative. (ii) Almost the entire population of CD2-positive cells turned Cerulean-positive upon induction. This correlation suggested that sorting for the CD2-positive cells would concomitantly enrich cells that are inducible and express Cerulean as well as the TALENs.

In a subsequent experiment I could increase the induction rate from 30 % to 84 % (Fig. 5.6) by sorting the cells for CD2 expression. This finding did not solve the problem of decreasing inducibility over time but it provided the possibility to keep the cells in a highly inducible state by continuous sorting for CD2 expression. It appears that the 5252 cell population is heterogeneous. Most likely the p5252 plasmid is under epigenetic control and cells are selected that do not express CD2 and have concomitantly lost their dox inducibility. Epigenetic repression is a frequent encounter and might indicate that either CD2 expression or a leaky expression of TALENs even in the absence of dox is not well tolerated by the cells. As a consequence selection of cells with an epigenetically repressed plasmid leads to its dox unresponsiveness.

Yet another important parameter was the rate of induction over time after adding dox to the system. 5252 and 5257 cells were induced and induction rates were continuously assessed the following five days. The mCherry expressing 5257 cells revealed an induction rate of 88 % on day one that constantly increased to almost 100 % on day five. In contrast, TALEN expressing 5252 cells showed a maximum induction rate of 43 % on day one dropping to 2 % on day five. This was a first hint that cells do not tolerate a high TALEN expression over a long period of time (Fig. 5.7). The issue of TALEN-mediated genotoxic effects is further addressed in 6.3.3.

In the first infection experiment I assessed the knockout of viral GFP at the protein level after infection with MHV-68-gfp. In HEK293 cells, 5257 cells (+/- dox) and 5252 cells (-dox) about 60 – 70 % of the cells turned GFP-positive after infection. In TALEN

expressing 5252 cells only 14 % of the cells were GFP-positive; a clear reduction compared to all other samples, indicating a TALEN-mediated knockout of the viral *gfp* sequence. Focusing on the Cerulean-positive fraction after infection revealed another interesting aspect (Fig. 5.8). The mean fluorescence intensity of Cerulean and therefore TALEN expression inversely correlates with the mean fluorescence intensity of GFP indicating that a high expression of TALEN 7 and 8 and the knockout of GFP correlate directly.

In the next experiment cleavage of the viral *gfp* sequence was analyzed at the level of DNA. By qPCR I quantified the *gfp* coding sequence versus the ORF6 sequence, which encodes an essential viral gene and therefore serves as a normalization control. Both genes, *gfp* and ORF6 are unique in the viral genome. Therefore, a ratio of one indicates the integrity of the *gfp* sequence, which decreases when a DSB is induced within the *gfp* sequence. qPCRs with uninduced 5252 and induced or uninduced 5257 cells revealed ratios of one as expected. The dox-mediated induction of TALEN expression in 5252 cells led to a ratio of 0.38 indicating that the viral *gfp* sequence was cleaved upon infection in the majority of cells (Fig. 5.9). In this PCR based assay the effect of TALEN expression was less pronounced as compared to the knockout at protein level measured by flow cytometry. This observation could reflect a repair of the TALEN-mediated DNA DSBs by the error prone NHEJ repair pathway. In this situation a mutation in the *gfp* coding sequence would lead to a knockout of the GFP protein but would not affect its suitability as a template in qPCR analysis.

### 6.3.3 Inhibition of viral replication and genotoxic effects

The first two infection experiments clearly showed a TALEN-mediated induction of DNA DSBs in the viral *gfp* sequence at the level of DNA and protein. In the third experiment I addressed the key question whether cleavage within the viral *gfp* sequence causes a reduction of viral replication. To address this question, 5252 and 5257 cells were infected with MHV-68-*gfp* with or without dox. The induction of mCherry expression in 5257 cells had no consequence for viral *de novo* synthesis whereas the induction of TALEN 7 and 8 expression in 5252 cells led to a reduction of viral *de novo* synthesis by a

factor of six indicating a clear antiviral effect of TALEN 7 and 8 expression (Fig. 5.10). This result has to be considered in relation to the induction rate of 85 %. In this particular experiment, 15 % of the cells (Cerulean-negative) likely supported viral replication limiting the theoretically achievable reduction to a factor of 6.7. Consequently, I tried to sort the cells after dox-mediated induction for Cerulean expression prior to viral infection. The fourth experiment in this series was independently performed two times with an induction rate of 5252 cells of 99 and 96 %, which led to a reduction of viral *de novo* synthesis by a factor of 67 and 71, respectively (Fig. 5.11).

Next, we asked if these reductions were solely due to a specific TALEN-mediated cleavage of the viral genome or a global effect of TALEN-mediated cytotoxicity. The latter would result in stressed cells, that might not support viral replication. To address this question, I infected 5252 cells with two different viruses. MHV-68-gfp was already used in the previous experiments revealing a TALEN-mediated reduction of viral replication. As a control I used a close relative of MHV-68-gfp. MHV-68 lacks the *gfp* coding sequence and thus the target sequence of the TALENs. Therefore, MHV-68 should replicate unaffected in TALEN expressing cells. This experiment was independently performed three times and in experiment one and three a reduction of MHV-68-gfp replication by a factor of 60 and 12 was shown. Although 92 % of the cells were Cerulean positive in the second experiment virus *de novo* synthesis was reduced by a factor of 9, only. A lower expression level of the TALENs indicated by the lower mean fluorescence intensity of Cerulean could have contributed to this observation in experiment 2 as compared to the higher GFP knockout rate in cells that express Cerulean at higher MFI as observed in Figure 5.8.

The infection of induced 5252 cells with the control MHV-68 virus revealed a 9-fold reduction in the first and third experiment at late time points clearly indicating an unspecific TALEN-mediated effect (Fig. 5.12). This effect was not observed in the second experiment suggesting that both reduction of viral replication and cytotoxic effects correlate with the expression levels of the TALENs.

TALEN-mediated cytotoxic effects, as indicated by two experiments (Fig. 5.7 and Fig. 5.12), compromise further applications in living organisms. The focus was therefore to



improve the functions of the TALENs in this setting. I generated three TALEN pairs based on TALEN 7 and 8 with mutated FokI domains (Fig. 5.13). The natural FokI enzyme confers homodimerization. Introducing different point mutations within the dimerization domain of FokI creates an asymmetric dimer interface due to electrostatic repulsion; a concept that has been proven successful with ZFNs. It was shown that these mutated FokI domains increase cleavage specificity and hence reduce unspecific effects and with it cytotoxicity (MILLER *et al.* 2007; SZCZEPEK *et al.* 2007; DOYON *et al.* 2011).

Previous experiments have revealed that the cleavage activity of TALENs is a crucial parameter, which can be limiting. Therefore, different mutated TALEN pairs equipped with wild type or mutated FokI domains were compared. Both, *in vitro* and *in vivo* assays showed a reduced cleavage activity of all TALEN pairs with mutated FokI domains compared to TALENs with wild type FokI domains. Among the mutated TALENs, the TALEN 7-ELD and TALEN 8-KKR pair showed the highest cleavage activity (Fig. 5.14 and Fig. 5.15). In addition, I assessed the cytotoxic potential of the different TALEN pairs. All TALEN pairs showed a reduced cytotoxicity compared to ZFNs, which are known to exhibit considerable ectopic cleavage activity. The experiment with mutated TALENs, however was not informative because the signals monitoring cytotoxicity were too close or within the range of the negative control (Fig. 5.16).

The TALEN-mediated genotoxic effect was also analyzed in the wild type TALENs expressing 5252 cell line. The cell line was induced with different concentrations of dox and the DNA damage response (DDR) signals were assessed at different time points post induction. The assay revealed a clear DDR signal at any condition tested, but a dose-dependent effect could not be detected (Fig. 5.17). On the basis of the previous experiments, we decided to test whether TALEN 7-ELD and TALEN 8-KKR reduce or block viral replication in combination with a reduced genotoxic effect.

As a prerequisite for this experiment, I established yet another inducible cell line (5480), which expresses TALEN 7-ELD, TALEN 8-KKR and Cerulean upon induction (Fig. 5.18). The cell line was infected with MHV-68-gfp and the *gfp* negative MHV-68 as a control. Infections with MHV-68 revealed no difference in viral *de novo* synthesis in uninduced and induced TALEN expressing cells as hoped. This observation suggested a reduced genotoxic effect compared to the wild type TALENs as seen in the experiments shown in

Figure 5.12. Infections with MHV-68-gfp revealed no reduction of viral replication in the first experiment and a weak effect in the second experiment, only. Previous results demonstrated that a high level of TALEN expression is crucial for an efficient cleavage of viral genomes in infected cells (Fig. 5.8, 5.11 and 5.12). The reduced cleavage activity of TALENs with mutated FokI domains does not considerably diminish the viral *de novo* synthesis presumably due to their reduced cleavage activity. This conclusion is in line with the observation that the slight reduction of viral replication in the second experiment correlates with a higher mean fluorescence intensity of Cerulean and therefore a higher expression of TALEN 7-ELD and 8-KKR as compared to the first experiment (Fig. 5.19).

A possible solution of this problem would be the highly active FokI sharkey domain (Guo *et al.* 2010) in combination with the ELD/KKR mutations. In this combination high locus-specificity should be conferred by the ELD/KKR mutations while maintaining a high nuclease activity due to the sharkey mutations.

In order to further illuminate the consequence of TALEN expression in the cells I performed an experiment assessing cellular proliferation and the induction rate of the different cell lines (Fig. 5.21). Prior to this experiment I established two new inducible cell lines, which served as controls in the following experiment. The previously used 5252 cell line expresses TALEN 7, TALEN 8 and Cerulean upon dox-mediated induction. The 5567 cell line expresses Cerulean, only, upon induction allowing us to assess the impact of Cerulean expression on cellular proliferation independent of TALEN 7 and 8 expression. The cell line 5568 expresses TALEN 7 and TALEN 8 similar to the cell line 5252 but Cerulean was replaced by the NGF-R surface marker. This cell line should show the same phenotype as the 5252 cell line excluding any effects arising from an interference between TALEN 7 and TALEN 8 with Cerulean or its coding sequence. Therefore, any phenotype in 5252 and 5568 as compared to 5567 would clearly be due to the expression of TALEN 7 and 8 (Fig. 5.20). The cellular proliferation of five different inducible cell lines was assessed with or without dox-mediated induction of protein expression. Two cell lines expressed TALEN 7 and 8 in combination with the different phenotypical markers Cerulean (5252) or NGF-R (5568). The cell line 5480 expresses the nucleases TALEN 7-ELD and TALEN 8-KKR and the phenotypical marker Cerulean.

Two control cell lines did not express any nucleases but the phenotypical markers mCherry (5257) or Cerulean (5567), only.

The experiment revealed that the expression of the phenotypical markers mCherry or Cerulean has little effect on cellular proliferation. The expression of the TALENs with wild type FokI domains led to a dramatic decrease of proliferation whereas the TALENs with mutated FokI domains showed an intermediate phenotype. This result documented a reduced genotoxicity of TALENs with mutated FokI domains compared to parental TALENs with wild type FokI domains. In parallel, the induction rate of the different cell lines was assessed. The TALEN 7 and 8 expressing cell lines (5252 and 5568) showed a maximum of expression 24 hours after dox-mediated induction constantly decreasing during the course of the experiment. This result suggests that the cells do not tolerate a long-lasting high expression of the TALENs, presumably due to their cytotoxic effect. The induced expression of the phenotypical markers (5257 and 5567) increased until day four reaching almost 100 % of the cell population. Again, the TALEN 7-ELD and TALEN 8-KKR expressing cell line (5480) showed an intermediate phenotype with a constant rate of marker expression from day one to three and a slight reduction on day four suggesting that the expression of these TALENs is tolerated presumably because they are less toxic to the cells.

The comparison with the phenotypical marker expressing cell lines suggested that the mutated TALENs can still cause a considerable level of cytotoxicity. Although TALENs are superior to ZFNs regarding this side effect HEK293 cells do not tolerate the expression of the nucleases longterm. Since the locus specificity of artificially generated nucleases is a crucial parameter, independent of their application, further efforts must focus on reducing off-target effects and cytotoxicity.

## 6.4 Possible applications

This thesis establishes a new concept in fighting cellular pathogens with double-stranded DNA genomes. A prerequisite for a possible application of this system is the reduction of TALEN-mediated genotoxic effects because a lasting TALEN expression

does not seem compatible with the survival of a cell at the moment. Therefore, the system has to be refined and improved.

One possible future application among others could be to fight epizootic diseases caused by viral infections in cattle breeding. For example, the bovine herpesvirus 1 (BHV-1) is known to cause clinical syndromes such as rhinotracheitis, pustular vulvovaginitis and balanoposthitis, abortion, infertility, conjunctivitis and encephalitis (NANDI *et al.* 2009). Although different types of vaccines are available for immunoprophylaxis the BHV-1 triggered diseases cause huge economic losses around the world (BISWAS *et al.* 2013). BHV-1 is also one of the etiologic factors causing the bovine respiratory disease, which generates an estimated loss of 640 million dollars in the United States annually (BOWLAND and SHEWEN 2000). Another example is the herpesvirus alcelaphine herpesvirus 1 (AIHV-1) and ovine herpesvirus 2 (OvHV-2), which cause inapparent infections in their reservoir host wildebeest and sheep, respectively, but upon infection of susceptible hosts including cattle, the viruses cause malignant catarrhal fever (RUSSELL *et al.* 2009). The disease is usually restricted to single animals but larger outbreaks were also reported (VAN WUIJCKHUISE-SJOUKE and KNIBBE 2007). Outbreaks in susceptible hosts result in a high mortality rate of 50 % about 50 days after exposure to an infected reservoir host (LI *et al.* 2006). Breeding herds of cattle, which are resistant to specific viruses are therefore a worthwhile goal. It would be conceivable to express the nucleases under control of a promoter, which is transactivated by a viral transcription factor, only. This would provide a viral trigger restricting TALEN expression to periods of infection and the infected cell, exclusively. Therefore, the organism would be intrinsically resistant to the virus but the detrimental effects of a long-term TALEN expression could be avoided.

Recently, a hypothesis was published suggesting that a specific bovine factor is an important contributor to the development of human colorectal cancer upon consumption of raw beef. One or more thermoresistant, potentially oncogenic and so far unknown bovine viruses (e.g., polyoma-, papilloma- or single-stranded DNA viruses) could lead to a latent infection in the colorectal tract increasing the risk for colorectal cancer (ZUR HAUSEN 2012). If this hypothesis proved true it would be appealing to generate virus-free herds of cattle, which do no longer propagate these oncogenic

viruses. This goal could be achieved by adopting anti-viral TALENs once the pathogens are identified.

## 6.5 Open questions and outlook

This thesis tested the possibility to lower or even block viral replication in infected cells by cleaving the viral genome with the aid of custom-made nucleases. A proof of principle was made by targeting the *gfp* coding sequence within the viral genome. Instead of using *gfp* as a model, a re-design of TALENs is an easy task in order to target genuine viral sequences and genes that are essential for viral propagation. The genome of large DNA viruses contains highly repetitive sequence arrays that constitute promising targets for antiviral TALENs. Besides these considerations there seem to be several parameters can be modified in order to improve the efficiency of the system.

A TALEN-mediated induction of a DNA double-strand break can be repaired by two cellular DNA repair pathways. Although NHEJ is an error prone repair mechanism, a repair of the DSB would result in an intact viral genome with a mutation within the *gfp* coding sequence. This would not affect viral *de novo* synthesis since GFP is dispensable for the virus. Furthermore, NHEJ could result in the loss of the TALENs cognate sequence within the viral genome resulting in a virus resistant to TALEN-mediated cleavage. Thus, the new mutant virus would replicate unhindered. This problem can be prevented by targeting an essential viral gene. Repair by NHEJ would likely result in a knockout of the targeted gene and therefore render the virus replication-incompetent. Hence, using TALENs targeting an essential viral gene would further increase the efficiency of the system since a single cut per viral genome would suffice to prevent viral *de novo* synthesis. An additional possibility is to target two viral genes simultaneously. First, this would result in a double knockout of two essential genes further preventing viral replication. Additionally, a double-cleavage would split the viral genome in several fragments. Consequently, the spatial distance between the ends of the DNA strands would impair the efficient DNA repair of the DSB.

In the course of this thesis it became also apparent that the cleavage efficiency of the nucleases is a major determinant. Therefore, the efficiency of the system could further be improved by increasing the cleavage activity of the TALENs. This could be done by using the highly efficient sharkey FokI domain (Guo *et al.* 2010) or by optimizing the TALENs' target sequence regarding their cleavage activity. This finding would also be another advantage when two viral genes are targeted with two TALEN pairs at the same time, since using two nuclease pairs should further increase the overall cleavage activity.

The system showed promising results but my experiments also revealed that the TALEN-mediated genotoxicity is a major drawback of the system, preventing an application in the near future. Therefore, further effort has to be put into generating highly specific TALENs in order to minimize cleavage at off-target sites. A first step in this direction was the generation of obligatory heterodimerizing FokI domains (Miller *et al.* 2007; SZCZEPEK *et al.* 2007; DOYON *et al.* 2011). Another approach was to combine zinc finger binding modules with the restriction enzyme PvuII. This strategy increases cleavage specificity since PvuII adds an extra element of specificity but restricts possible target sites to given PvuII sites with adjacent zinc finger binding sites (SCHIERLING *et al.* 2012). An additional parameter would be to optimize the DNA binding domain in order to increase the specificity of the TALENs or extend the recognized motif sequence of the TALENs. A systematic approach revealed that binding affinity and transcriptional activation of TALEs strongly depends on the RVD composition and on the target sequence context (MECKLER *et al.* 2013). Yet, it remains uncertain if a high locus specificity, which allows a constitutive expression of the TALENs is achievable. This concern is based on the relatively weak dissociation constant of TALENs (Fig. 4.6) and their presumed "sliding" on DNA while scanning for their cognate binding motifs. There is clearly room for improvement but it is currently unclear, whether all problems can be solved to perfection. In 2011 it was shown that the expression of virus-specific meganucleases resulted in the cleavage of the viral genome of the DNA virus HSV1 in the expressing cells upon infection.

The cleavage of the viral genome resulted in a reduced viral load, whereas the TALEN-based system shown here seems to prevent viral *de novo* synthesis more efficient as compared to the meganuclease-based system. Another advantage of TALENs is their simple engineering and the high flexibility allowing to target virtually all sequences,

whereas the engineering of meganucleases with novel target sequences remains restricted (GROSSE *et al.* 2011).

Recently, a new technology in genomic engineering with endonucleases was reported. The novel system is based on the type II clustered, regularly interspaced, short palindromic repeats (CRISPR) interference system, which is part of adaptive immunity in bacteria and archaea. The system consists of the Cas9 endonuclease and two short RNA molecules, CRISPR RNA (crRNA) and transactivating crRNA (tracrRNA). The endonuclease forms a complex with the two RNA molecules, which guide the enzyme to the target site, where cleavage occurs. In bacteria, the system targets the genome of bacteriophages as a part of immune defence. It was shown that the two RNAs can be replaced by the chimeric single guide RNA (sgRNA). Therefore, the nuclease system can be easily directed to new target sites by modifying the sgRNA sequence (JINEK *et al.* 2013). Cas9 was already used to modify the genome of zebrafish, mouse and human cells with a efficiency comparable to ZFNs and TALENs (CONG *et al.* 2013; HWANG *et al.* 2013; MALI *et al.* 2013). The relatively short target sequence of 20 bp and the fact that mismatches are tolerated at the 5' end of the target sequence (JIANG *et al.* 2013) raises the question whether RNA-guided nucleases (RGEN) have a better locus specificity. Several groups have not observed overt cytotoxicity (CHO *et al.* 2013; CONG *et al.* 2013; JIANG *et al.* 2013; JINEK *et al.* 2013; MALI *et al.* 2013) but one group observed cytotoxicity in zebrafish comparable to ZFNs and TALENs. Differently composed sgRNAs seem to show different levels of cytotoxicity (HWANG *et al.* 2013). These are first promising results but it is also attractive to further illuminate the target site specificity and investigate whether RGENs can be modified to recognize longer target sites (MUSSOLINO and CATHOMEN 2013). This newly developed system appears to be an interesting alternative to TALENs in the construction of an antiviral system, and could extend the perspective frame of my experimental work.

## 7. Summary

Custom-made nucleases with predetermined target sites, like zinc finger nucleases (ZFNs) and TALE nucleases (TALENs), facilitated precise modifications of the genomes of cells and transgenic organisms. The locus-specific induction of DNA double-strand breaks (DSBs) exploits the cellular DNA repair pathways that mediate homologous recombination and non-homologous DNA end-joining. This so-called genomic engineering can be used to achieve a knockout, deletion or correction of a given gene or even add novel genes and genetic elements. The aim of this thesis was to establish a novel application: a TALEN-based system that prevents viral replication by inducing a DSB within the viral genome.

A set of novel TALENs were constructed, which target a DNA sequence motif within the GFP encoding gene. The newly generated TALENs show efficient cleavage in *in vitro* and *in vivo* assays. Inducible cell lines, which permit the conditional expression of the nucleases were established. Infection of TALEN expressing cells with the *gfp* containing herpes virus MHV-68-*gfp* showed a dramatically reduced viral *de novo* synthesis but TALEN expression can have also unspecific effects on the infected cells and virus synthesis. TALENs act as homodimers but mutated FokI cleavage domains, which force heterodimerization of the TALENs can reduce unwanted off-target effects as shown with ZFNs. Heterodimeric TALENs that encompass mutated FokI domains showed reduced cytotoxicity but also revealed a lower cleavage activity that did not suffice to efficiently inhibit viral replication.

The long-term effect of TALEN expression on the cells was studied by assessing cellular proliferation as a function of TALEN expression. TALEN expressing cells showed a considerable reduction of proliferation as compared to control cell lines expressing phenotypical markers, only, indicative of cytotoxic effects. The cell line expressing the TALENs with mutated FokI domains showed an intermediate phenotype suggesting a



reduced but still considerable cytotoxic potential of the TALENs with mutated FokI domains. Cytotoxicity of TALENs could be attributed at least in part to a genotoxic effect as the cellular DNA damage response pathway was activated upon TALEN expression.

In a biochemical approach, the specificity of DNA binding and the dissociation constants of ZFNs and TALEs were analyzed. ZF $\Delta$ N (ZFN with inactivated nuclease domain) showed a relatively low binding affinity with a  $K_d$  value of around 300 nM, independent of the size of the DNA template.  $K_d$  values for ZF $\Delta$ N dimers could not be assessed due to a low stability of the complex. A synthetic TALE showed a  $K_d$  value of 300 nM only with a longer DNA template whereas binding to a short DNA template was too weak to determine the  $K_d$  value. This finding indicates that the TALE DNA binding affinity is a function of the length of the DNA template. Experiments with the natural TALE HAX3 using a short and a longer template delivered  $K_d$  values of 674 nM and 76 nM, respectively, confirming the higher binding affinities of TALEs towards longer DNA templates.

In summary, the results of this thesis revealed that TALENs can serve as suitable antiviral agents. High level expression of TALENs, which is necessary to achieve an efficient reduction of viral *de novo* synthesis, causes cytotoxicity at a level, which is not compatible with cellular survival. Therefore, further research is necessary to reduce off-target effects of these custom-made nucleases improving the TALEN-based antiviral system.

## 8. Abbreviations

$\alpha$	anti (in context of antibody specificity)
A	adenosine
aa	amino acid
AIHV-1	alceplaphine herpesvirus 1
AP-1	activator protein 1
APS	ammonium persulfate
BAC	bacterial artificial chromosome
BHK21	baby hamster kidney 21
BHV-1	bovine herpesvirus 1
bp	base pair
BSA	bovine serum albumin
C	cytosine
C terminus	carboxy terminus
CMV	cytomegalovirus
Cpm	counts per minute
CRISPR	clustered, regulatory interspaced, short palindromic repeats
crRNA	CRISPR RNA
D	Aspartic acid
DDR	DNA damage response
dox	doxycycline
DMEM	Dulbecco's Modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSB	double-strand break
dsDNA	double-stranded DNA
<i>E.coli</i>	Escherichia coli
EBV	Epstein-Barr virus

EDTA	ethyldiamintetraacetic acid
EMSA	electrophoretic mobility shift assay
FACS	fluorescent-activated cell sorting
FCS	fetal calf serum
Fig.	figure
FITC	fluorescein isothiocyanate
FSC	forward scatter
G	guanosine
<i>gfp</i>	GFP encoding sequence
GFP	green fluorescent protein
H	Histidine
HEK293	human embryonic kidney 293
HIV	human immunodeficiency virus
HR	homologous recombination
HRP	horseradish peroxidase
I	Isoleucine
K	Lysine
Kb	kilo base pair
K <sub>d</sub>	equilibrium dissociation constant
kDa	kilo Dalton
KSHV	Kaposi Sarcoma associated virus
LB	Luria broth
MACS	magnetic activated cell sorting
MCS	multiple cloning site
MHV-68	murine herpesvirus 68
MHV-68- <i>gfp</i>	murine herpesvirus 68 harboring the <i>gfp</i> gene
MOI	multiplicity of infections
N	Asparagine
NGF-R	nerve growth factor receptor
NLS	nuclear localization signal
NHEJ	non-homologous DNA end joining
N terminus	amino terminus
oriP	origin of plasmid replication
OvHV-2	ovine herpesvirus 2

PAA	polyacryamide
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEI	polyethylenimin
pH2A.X, $\gamma$ H2A.X	phosphorylated histone H2A.X
qPCR	quantitative PCR
RGEN	RNA-guided nuclease
RNA	ribonucleic acid
RPMI	Rosewell Park Memorial Institute
RVD	repeat variable di-residue
S	Serine
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SDSA	synthesis dependent strand annealing
sgRNA	single guide RNA
SSC	sideward scatter
ssDNA	single strand DNA
T	thymine
TAE	buffer containing Tris base, acetic acid and EDTA
TALE	transcription activator like effector
TALEN	TALE nuclease
TBE	buffer containing Tris base, boric acid and EDTA
TEMED	tetramethylethylenediamine
tracrRNA	transactivating crRNA
Tris	trishydroxymethylaminomethane
WB	Western blot
ZFN	zinc finger nuclease
ZF $\Delta$ N	ZFN with inactivated nuclease domain

## 9. Literature

- ADLER, H., M. MESSERLE and U. H. KOSZINOWSKI, 2001 Virus reconstituted from infectious bacterial artificial chromosome (BAC)-cloned murine gammaherpesvirus 68 acquires wild-type properties in vivo only after excision of BAC vector sequences. *J Virol* **75**: 5692-5696.
- ADLER, H., M. MESSERLE, M. WAGNER and U. H. KOSZINOWSKI, 2000 Cloning and mutagenesis of the murine gammaherpesvirus 68 genome as an infectious bacterial artificial chromosome. *J Virol* **74**: 6964-6974.
- ALWIN, S., M. B. GERE, E. GUHL, K. EFFERTZ, C. F. BARBAS, 3RD *et al.*, 2005 Custom zinc-finger nucleases for use in human cells. *Mol Ther* **12**: 610-617.
- BEERLI, R. R., D. J. SEGAL, B. DREIER and C. F. BARBAS, 3RD, 1998 Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci U S A* **95**: 14628-14633.
- BELKHADIR, Y., R. SUBRAMANIAM and J. L. DANGL, 2004 Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Curr Opin Plant Biol* **7**: 391-399.
- BEUMER, K., G. BHATTACHARYYA, M. BIBIKOVA, J. K. TRAUTMAN and D. CARROLL, 2006 Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics* **172**: 2391-2403.
- BHAKTA, M. S., I. M. HENRY, D. G. OUSTEROUT, K. T. DAS, S. H. LOCKWOOD *et al.*, 2013 Highly active zinc-finger nucleases by extended modular assembly. *Genome Res* **23**: 530-538.
- BISWAS, S., S. BANDYOPADHYAY, U. DIMRI and H. P. P, 2013 Bovine herpesvirus-1 (BHV-1) - a re-emerging concern in livestock: a revisit to its biology, epidemiology, diagnosis, and prophylaxis. *Vet Q* **33**: 68-81.
- BITINAITE, J., D. A. WAH, A. K. AGGARWAL and I. SCHILDKRAUT, 1998 FokI dimerization is required for DNA cleavage. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 10570-10575.
- BLANCAFORT, P., D. J. SEGAL and C. F. BARBAS, 3RD, 2004 Designing transcription factor architectures for drug discovery. *Mol Pharmacol* **66**: 1361-1371.
- BOCH, J., H. SCHOLZE, S. SCHORNACK, A. LANDGRAF, S. HAHN *et al.*, 2009 Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**: 1509-1512.
- BOGDANOVE, A. J., S. SCHORNACK and T. LAHAYE, 2010 TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol* **13**: 394-401.
- BOLLER, T., and S. Y. HE, 2009 Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science* **324**: 742-744.
- BONAS, U., R. E. STALL and B. STASKAWICZ, 1989 Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet* **218**: 127-136.

- BORNKAMM, G. W., C. BERENS, C. KUKLIK-ROOS, J. M. BECHET, G. LAUX *et al.*, 2005 Stringent doxycycline-dependent control of gene activities using an episomal one-vector system. *Nucleic Acids Res* **33**: e137.
- BOWLAND, S. L., and P. E. SHEWEN, 2000 Bovine respiratory disease: commercial vaccines currently available in Canada. *Can Vet J* **41**: 33-48.
- BROWN, R. S., C. SANDER and P. ARGOS, 1985 The primary structure of transcription factor TFIIIA has 12 consecutive repeats. *FEBS Lett* **186**: 271-274.
- BULTMANN, S., R. MORBITZER, C. S. SCHMIDT, K. THANISCH, F. SPADA *et al.*, 2012 Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res* **40**: 5368-5377.
- BUSHMAN, F. D., and M. D. MILLER, 1997 Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *J Virol* **71**: 458-464.
- CATHOMEN, T., and J. K. JOUNG, 2008 Zinc-finger nucleases: the next generation emerges. *Mol Ther* **16**: 1200-1207.
- CATTO, L. E., S. GANGULY, S. E. MILSON, A. J. WELSH and S. E. HALFORD, 2006 Protein assembly and DNA looping by the FokI restriction endonuclease. *Nucleic Acids Res* **34**: 1711-1720.
- CERMAK, T., E. L. DOYLE, M. CHRISTIAN, L. WANG, Y. ZHANG *et al.*, 2011 Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**: e82.
- CHO, S. W., S. KIM, J. M. KIM and J. S. KIM, 2013 Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**: 230-232.
- CHRISTIAN, M., T. CERMAK, E. L. DOYLE, C. SCHMIDT, F. ZHANG *et al.*, 2010 Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**: 757-761.
- CHRISTIAN, M. L., Z. L. DEMOREST, C. G. STARKER, M. J. OSBORN, M. D. NYQUIST *et al.*, 2012 Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. *PLoS One* **7**: e45383.
- CHYTIL, M., B. R. PETERSON, D. A. ERLANSON and G. L. VERDINE, 1998 The orientation of the AP-1 heterodimer on DNA strongly affects transcriptional potency. *Proc Natl Acad Sci U S A* **95**: 14076-14081.
- CONG, L., F. A. RAN, D. COX, S. L. LIN, R. BARRETTO *et al.*, 2013 Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**: 819-823.
- CORNU, T. I., S. THIBODEAU-BEGANNY, E. GUHL, S. ALWIN, M. EICHTINGER *et al.*, 2008 DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther* **16**: 352-358.
- DAHLEM, T. J., K. HOSHIJIMA, M. J. JURYNEC, D. GUNTHER, C. G. STARKER *et al.*, 2012 Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genetics* **8**: e1002861.
- DELECLUSE, H. J., T. HILSENDEGEN, D. PICH, R. ZEIDLER and W. HAMMERSCHMIDT, 1998 Propagation and recovery of intact, infectious Epstein-Barr virus from prokaryotic to human cells. *Proc Natl Acad Sci U S A* **95**: 8245-8250.
- DENG, D., C. YAN, X. PAN, M. MAHFOUZ, J. WANG *et al.*, 2012 Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335**: 720-723.
- DESJARLAIS, J. R., and J. M. BERG, 1993 Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. *Proc Natl Acad Sci U S A* **90**: 2256-2260.
- DOYON, Y., T. D. VO, M. C. MENDEL, S. G. GREENBERG, J. WANG *et al.*, 2011 Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods* **8**: 74-79.

- ELROD-ERICKSON, M., M. A. ROULD, L. NEKLUDOVA and C. O. PABO, 1996 Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure* **4**: 1171-1180.
- EPSTEIN, M. A., B. G. ACHONG and Y. M. BARR, 1964 Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. *Lancet* **1**: 702-703.
- FRANKEL, A. D., J. M. BERG and C. O. PABO, 1987 Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc Natl Acad Sci U S A* **84**: 4841-4845.
- FRIED, M., and D. M. CROTHERS, 1981 Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res* **9**: 6505-6525.
- GAJ, T., C. A. GERSBACH and C. F. BARBAS, 3RD, 2013 ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.*
- GAO, H., X. WU, J. CHAI and Z. HAN, 2012 Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. *Cell Res* **22**: 1716-1720.
- GLOECKNER, C. J., K. BOLDT, A. SCHUMACHER, R. ROEPMAN and M. UEFFING, 2007 A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics* **7**: 4228-4234.
- GOHRE, V., and S. ROBATZEK, 2008 Breaking the barriers: microbial effector molecules subvert plant immunity. *Annu Rev Phytopathol* **46**: 189-215.
- GRAHAM, F. L., J. SMILEY, W. C. RUSSELL and R. NAIRN, 1977 Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol* **36**: 59-74.
- GRAWUNDER, U., M. WILM, X. WU, P. KULESZA, T. E. WILSON *et al.*, 1997 Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388**: 492-495.
- GROSSE, S., N. HUOT, C. MAHIET, S. ARNOULD, S. BARRADEAU *et al.*, 2011 Meganuclease-mediated Inhibition of HSV1 Infection in Cultured Cells. *Mol Ther* **19**: 694-702.
- GU, J., H. LU, A. G. TSAI, K. SCHWARZ and M. R. LIEBER, 2007 Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence. *Nucleic Acids Res* **35**: 5755-5762.
- GUO, J., T. GAJ and C. F. BARBAS, 3RD, 2010 Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J Mol Biol* **400**: 96-107.
- HAMMAR, P., P. LEROY, A. MAHMUTOVIC, E. G. MARKLUND, O. G. BERG *et al.*, 2012 The lac repressor displays facilitated diffusion in living cells. *Science* **336**: 1595-1598.
- HANAHAN, D., 1985 Techniques for transformation of *E. coli*. In Glover, D. (ed.), *DNA Cloning: A Practical Approach*. **1**: 109-135.
- HANDEL, E. M., S. ALWIN and T. CATHOMEN, 2009 Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol Ther* **17**: 104-111.
- HEYER, W. D., K. T. EHMSSEN and J. LIU, 2010 Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* **44**: 113-139.
- HIGUCHI, R., C. FOCKLER, G. DOLLINGER and R. WATSON, 1993 Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N Y)* **11**: 1026-1030.
- HOCKEMEYER, D., H. WANG, S. KIANI, C. S. LAI, Q. GAO *et al.*, 2011 Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**: 731-734.
- HOEIJMAKERS, J. H., 2001 Genome maintenance mechanisms for preventing cancer. *Nature* **411**: 366-374.
- HUANG, P., A. XIAO, M. G. ZHOU, Z. Y. ZHU, S. LIN *et al.*, 2011 Heritable gene targeting in zebrafish using customized TALENs. *Nature Biotechnology* **29**: 699-700.

- HURT, J. A., S. A. THIBODEAU, A. S. HIRSH, C. O. PABO and J. K. JOUNG, 2003 Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proc Natl Acad Sci U S A* **100**: 12271-12276.
- HWANG, W. Y., Y. FU, D. REYON, M. L. MAEDER, S. Q. TSAI *et al.*, 2013 Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**: 227-229.
- ISALAN, M., Y. CHOO and A. KLUG, 1997 Synergy between adjacent zinc fingers in sequence-specific DNA recognition. *Proc Natl Acad Sci U S A* **94**: 5617-5621.
- JACOBS, G. H., 1992 Determination of the base recognition positions of zinc fingers from sequence analysis. *EMBO J* **11**: 4507-4517.
- JAMIESON, A. C., S. H. KIM and J. A. WELLS, 1994 In vitro selection of zinc fingers with altered DNA-binding specificity. *Biochemistry* **33**: 5689-5695.
- JAMIESON, A. C., J. C. MILLER and C. O. PABO, 2003 Drug discovery with engineered zinc-finger proteins. *Nat Rev Drug Discov* **2**: 361-368.
- JIANG, W., D. BIKARD, D. COX, F. ZHANG and L. A. MARRAFFINI, 2013 RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* **31**: 233-239.
- JINEK, M., A. EAST, A. CHENG, S. LIN, E. MA *et al.*, 2013 RNA-programmed genome editing in human cells. *Elife* **2**: e00471.
- JOHN, M., R. LEPIK, S. J. BUSCH, M. GRANGER-SCHNARR and M. SCHNARR, 1996 DNA binding of Jun and Fos bZip domains: homodimers and heterodimers induce a DNA conformational change in solution. *Nucleic Acids Res* **24**: 4487-4494.
- JOHNSON, R. D., and M. JASIN, 2001 Double-strand-break-induced homologous recombination in mammalian cells. *Biochem Soc Trans* **29**: 196-201.
- JONES, J. D., and J. L. DANGL, 2006 The plant immune system. *Nature* **444**: 323-329.
- JOUNG, J. K., E. I. RAMM and C. O. PABO, 2000 A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A* **97**: 7382-7387.
- KAY, S., and U. BONAS, 2009 How *Xanthomonas* type III effectors manipulate the host plant. *Curr Opin Microbiol* **12**: 37-43.
- KAY, S., S. HAHN, E. MAROIS, G. HAUSE and U. BONAS, 2007 A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**: 648-651.
- KIM, H. J., H. J. LEE, H. KIM, S. W. CHO and J. S. KIM, 2009 Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res* **19**: 1279-1288.
- KIM, J. S., and C. O. PABO, 1998 Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. *Proc Natl Acad Sci U S A* **95**: 2812-2817.
- KIM, Y. G., J. CHA and S. CHANDRASEGARAN, 1996 Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**: 1156-1160.
- KREJCI, L., V. ALTMANNOVA, M. SPIREK and X. ZHAO, 2012 Homologous recombination and its regulation. *Nucleic Acids Res* **40**: 5795-5818.
- LI, H., N. TAUS, C. JONES, B. MURPHY, J. EVERMANN *et al.*, 2006 A devastating outbreak of malignant catarrhal fever in a bison feedlot. *J Vet Diagn Invest*.
- LI, L., and S. CHANDRASEGARAN, 1993 Alteration of the cleavage distance of Fok I restriction endonuclease by insertion mutagenesis. *Proc Natl Acad Sci U S A* **90**: 2764-2768.
- LI, L., L. P. WU and S. CHANDRASEGARAN, 1992 Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci U S A* **89**: 4275-4279.
- LI, T., S. HUANG, W. Z. JIANG, D. WRIGHT, M. H. SPALDING *et al.*, 2011 TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res* **39**: 359-372.



- LI, X., and W. D. HEYER, 2008 Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* **18**: 99-113.
- LIEBER, M. R., 1999 The biochemistry and biological significance of nonhomologous DNA end joining: an essential repair process in multicellular eukaryotes. *Genes Cells* **4**: 77-85.
- LIEBER, M. R., 2008 The mechanism of human nonhomologous DNA end joining. *J Biol Chem* **283**: 1-5.
- LIEBER, M. R., 2010 The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**: 181-211.
- LIEBER, M. R., and Z. E. KARANJAWALA, 2004 Ageing, repetitive genomes and DNA damage. *Nat Rev Mol Cell Biol* **5**: 69-75.
- LIEBER, M. R., Y. MA, U. PANNICKE and K. SCHWARZ, 2003 Mechanism and regulation of human non-homologous DNA end-joining. *Nat Rev Mol Cell Biol* **4**: 712-720.
- LIEBER, M. R., Y. MA, U. PANNICKE and K. SCHWARZ, 2004 The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. *DNA Repair (Amst)* **3**: 817-826.
- LIU, J., C. LI, Z. YU, P. HUANG, H. WU *et al.*, 2012 Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. *J Genet Genomics* **39**: 209-215.
- LIU, Q., D. J. SEGAL, J. B. GHIARA and C. F. BARBAS, 3RD, 1997 Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc Natl Acad Sci U S A* **94**: 5525-5530.
- LOMBARDO, A., P. GENOVESE, C. M. BEAUSEJOUR, S. COLLEONI, Y. L. LEE *et al.*, 2007 Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* **25**: 1298-1306.
- LOZZIO, C. B., and B. B. LOZZIO, 1975 Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood* **45**: 321-334.
- MA, Y., H. LU, B. TIPPIN, M. F. GOODMAN, N. SHIMAZAKI *et al.*, 2004 A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol Cell* **16**: 701-713.
- MA, Y., U. PANNICKE, K. SCHWARZ and M. R. LIEBER, 2002 Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108**: 781-794.
- MA, Y., K. SCHWARZ and M. R. LIEBER, 2005 The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amst)* **4**: 845-851.
- MACKAY, J. P., and M. CROSSLEY, 1998 Zinc fingers are sticking together. *Trends Biochem Sci* **23**: 1-4.
- MAEDER, M. L., S. THIBODEAU-BEGANNY, A. OSIAK, D. A. WRIGHT, R. M. ANTHONY *et al.*, 2008 Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* **31**: 294-301.
- MAHFOUZ, M. M., L. LI, M. SHAMIMUZZAMAN, A. WIBOWO, X. FANG *et al.*, 2011 De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc Natl Acad Sci U S A* **108**: 2623-2628.
- MAK, A. N., P. BRADLEY, R. A. CERNADAS, A. J. BOGDANOVE and B. L. STODDARD, 2012 The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335**: 716-719.
- MALI, P., L. YANG, K. M. ESVELT, J. AACH, M. GUELL *et al.*, 2013 RNA-guided human genome engineering via Cas9. *Science* **339**: 823-826.
- MALKOVA, A., E. L. IVANOV and J. E. HABER, 1996 Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proc Natl Acad Sci U S A* **93**: 7131-7136.

- MECKLER, J. F., M. S. BHAKTA, M. S. KIM, R. OVADIA, C. H. HABRIAN *et al.*, 2013 Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res* **41**: 4118-4128.
- MENG, X., M. B. NOYES, L. J. ZHU, N. D. LAWSON and S. A. WOLFE, 2008 Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol* **26**: 695-701.
- MILLER, J., A. D. McLACHLAN and A. KLUG, 1985 Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* **4**: 1609-1614.
- MILLER, J. C., M. C. HOLMES, J. WANG, D. Y. GUSCHIN, Y. L. LEE *et al.*, 2007 An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* **25**: 778-785.
- MILLER, J. C., S. TAN, G. QIAO, K. A. BARLOW, J. WANG *et al.*, 2011 A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**: 143-148.
- MIMITOU, E. P., and L. S. SYMINGTON, 2008 Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455**: 770-774.
- MORBITZER, R., J. ELSAESSER, J. HAUSNER and T. LAHAYE, 2011 Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res* **39**: 5790-5799.
- MORBITZER, R., P. ROMER, J. BOCH and T. LAHAYE, 2010 Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc Natl Acad Sci U S A* **107**: 21617-21622.
- MORRISON, C., and S. TAKEDA, 2000 Genetic analysis of homologous DNA recombination in vertebrate somatic cells. *Int J Biochem Cell Biol* **32**: 817-831.
- MOSCOU, M. J., and A. J. BOGDANOVE, 2009 A simple cipher governs DNA recognition by TAL effectors. *Science* **326**: 1501.
- MULLIS, K., F. FALOONA, S. SCHARF, R. SAIKI, G. HORN *et al.*, 1986 Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51 Pt 1**: 263-273.
- MUSSOLINO, C., and T. CATHOMEN, 2013 RNA guides genome engineering. *Nat Biotechnol* **31**: 208-209.
- MUSSOLINO, C., R. MORBITZER, F. LUTGE, N. DANNEMANN, T. LAHAYE *et al.*, 2011 A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res* **39**: 9283-9293.
- NANDI, S., M. KUMAR, M. MANOHAR and R. S. CHAUHAN, 2009 Bovine herpes virus infections in cattle. *Anim Health Res Rev* **10**: 85-98.
- NASSIF, N., J. PENNEY, S. PAL, W. R. ENGELS and G. B. GLOOR, 1994 Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* **14**: 1613-1625.
- PAQUES, F., and J. E. HABER, 1999 Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**: 349-404.
- PAVLETICH, N. P., and C. O. PABO, 1991 Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**: 809-817.
- PELLEGRINO, G. R., and J. M. BERG, 1991 Identification and characterization of "zinc-finger" domains by the polymerase chain reaction. *Proc Natl Acad Sci U S A* **88**: 671-675.
- PEREZ, E. E., J. WANG, J. C. MILLER, Y. JOUVENOT, K. A. KIM *et al.*, 2008 Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**: 808-816.
- POFFENBERGER, K. L., and B. ROIZMAN, 1985 A noninverting genome of a viable herpes simplex virus 1: presence of head-to-tail linkages in packaged genomes and requirements for circularization after infection. *J Virol* **53**: 587-595.

- PONTEN, J., and E. SAKSELA, 1967 Two established in vitro cell lines from human mesenchymal tumours. *Int J Cancer* **2**: 434-447.
- PORTEUS, M. H., and D. BALTIMORE, 2003 Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**: 763.
- PORTEUS, M. H., and D. CARROLL, 2005 Gene targeting using zinc finger nucleases. *Nat Biotechnol* **23**: 967-973.
- PORTEUS, M. H., T. CATHOMEN, M. D. WEITZMAN and D. BALTIMORE, 2003 Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Mol Cell Biol* **23**: 3558-3565.
- PRUETT-MILLER, S. M., J. P. CONNELLY, M. L. MAEDER, J. K. JOUNG and M. H. PORTEUS, 2008 Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. *Mol Ther* **16**: 707-717.
- RAMIREZ, C. L., J. E. FOLEY, D. A. WRIGHT, F. MULLER-LERCH, S. H. RAHMAN *et al.*, 2008 Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods* **5**: 374-375.
- REBAR, E. J., H. A. GREISMAN and C. O. PABO, 1996 Phage display methods for selecting zinc finger proteins with novel DNA-binding specificities. *Methods Enzymol* **267**: 129-149.
- REBAR, E. J., and C. O. PABO, 1994 Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **263**: 671-673.
- REYON, D., S. Q. TSAI, C. KHAYTER, J. A. FODEN, J. D. SANDER *et al.*, 2012 FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* **30**: 460-465.
- ROGAKOU, E. P., C. BOON, C. REDON and W. M. BONNER, 1999 Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J Cell Biol* **146**: 905-916.
- ROIZMAN, B., and P. E. PELLET, 2007 The family herpesviridae: A brief introduction. *Fields Virology*, 5th edition Vol. 2 (Chapter 66): 2479-2499.
- RUSSELL, G., J. STEQART and D. HAIG, 2009 Malignant catarrhal fever: a review. *Vet J* **179**: 324-335.
- RYDER, S. P., M. I. RECHT and J. R. WILLIAMSON, 2008 Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods Mol Biol* **488**: 99-115.
- SAMBROOK, J. A. R., D., 2001 *Molecular Cloning A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press **Third Edition**.
- SCHIERLING, B., N. DANNEMANN, L. GABSALILOW, W. WENDE, T. CATHOMEN *et al.*, 2012 A novel zinc-finger nuclease platform with a sequence-specific cleavage module. *Nucleic Acids Res* **40**: 2623-2638.
- SCHMID-BURCK, J. L., T. SCHMIDT, V. KAISER, K. HONING and V. HORNING, 2013 A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotechnol* **31**: 76-81.
- SCHORNACK, S., A. MEYER, P. ROMER, T. JORDAN and T. LAHAYE, 2006 Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. *J Plant Physiol* **163**: 256-272.
- SHASTRY, B. S., 1996 Transcription factor IIIA (TFIIIA) in the second decade. *J Cell Sci* **109**: 535-539.
- SHIMIZU, Y., C. SOLLU, J. F. MECKLER, A. ADRIAENSSENS, A. ZYKOVICH *et al.*, 2011 Adding fingers to an engineered zinc finger nuclease can reduce activity. *Biochemistry* **50**: 5033-5041.
- SMITH, A. E., and K. G. FORD, 2007 Specific targeting of cytosine methylation to DNA sequences in vivo. *Nucleic Acids Res* **35**: 740-754.

- SMITH, J., M. BIBIKOVA, F. G. WHITBY, A. R. REDDY, S. CHANDRASEGARAN *et al.*, 2000 Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res* **28**: 3361-3369.
- SOLLU, C., K. PARS, T. I. CORNU, S. THIBODEAU-BEGANNY, M. L. MAEDER *et al.*, 2010 Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids Res* **38**: 8269-8276.
- SONG, M. J., S. HWANG, W. H. WONG, T. T. WU, S. LEE *et al.*, 2005 Identification of viral genes essential for replication of murine gamma-herpesvirus 68 using signature-tagged mutagenesis. *Proc Natl Acad Sci U S A* **102**: 3805-3810.
- STOKER, M., and I. MACPHERSON, 1964 Syrian Hamster Fibroblast Cell Line Bhk21 and Its Derivatives. *Nature* **203**: 1355-1357.
- SZCZEPKE, M., V. BRONDANI, J. BUCHEL, L. SERRANO, D. J. SEGAL *et al.*, 2007 Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* **25**: 786-793.
- SZOSTAK, J. W., T. L. ORR-WEAVER, R. J. ROTHSTEIN and F. W. STAHL, 1983 The double-strand-break repair model for recombination. *Cell* **33**: 25-35.
- SZUREK, B., E. MAROIS, U. BONAS and G. VAN DEN ACKERVEKEN, 2001 Eukaryotic features of the *Xanthomonas* type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. *Plant J* **26**: 523-534.
- SZUREK, B., O. ROSSIER, G. HAUSE and U. BONAS, 2002 Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. *Mol Microbiol* **46**: 13-23.
- TAKATA, M., M. S. SASAKI, E. SONODA, C. MORRISON, M. HASHIMOTO *et al.*, 1998 Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* **17**: 5497-5508.
- TAYLOR, G. S., and D. J. BLACKBOURN, 2011 Infectious agents in human cancers: lessons in immunity and immunomodulation from gammaherpesviruses EBV and KSHV. *Cancer Lett* **305**: 263-278.
- TESSON, L., C. USAL, S. MENORET, E. LEUNG, B. J. NILES *et al.*, 2011 Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol* **29**: 695-696.
- TUPLER, R., G. PERINI and M. R. GREEN, 2001 Expressing the human genome. *Nature* **409**: 832-833.
- URNOV, F. D., J. C. MILLER, Y. L. LEE, C. M. BEAUSEJOUR, J. M. ROCK *et al.*, 2005 Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**: 646-651.
- VAN WUIJCKHUISE-SJOUKE, L., and G. KNIBBE, 2007 Large outbreak of malignant catarrhal fever in cattle. *Tijdschr Diergeneeskde* **132**: 732-734.
- VANAMEE, E. S., S. SANTAGATA and A. K. AGGARWAL, 2001 FokI requires two specific DNA sites for cleavage. *J Mol Biol* **309**: 69-78.
- VIRGIN, H. W. T., P. LATREILLE, P. WAMSLEY, K. HALLSWORTH, K. E. WECK *et al.*, 1997 Complete sequence and genomic analysis of murine gammaherpesvirus 68. *J Virol* **71**: 5894-5904.
- WAUGH, D. S., and R. T. SAUER, 1993 Single amino acid substitutions uncouple the DNA binding and strand scission activities of Fok I endonuclease. *Proc Natl Acad Sci U S A* **90**: 9596-9600.
- WOLFE, S. A., E. I. RAMM and C. O. PABO, 2000 Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers. *Structure* **8**: 739-750.

- WOODCOCK, D. M., P. J. CROWTHER, J. DOHERTY, S. JEFFERSON, E. DECROUZ *et al.*, 1989 Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res* **17**: 3469-3478.
- WRIGHT, D. A., J. A. TOWNSEND, R. J. WINFREY, JR., P. A. IRWIN, J. RAJAGOPAL *et al.*, 2005 High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J* **44**: 693-705.
- WU, T. T., L. TONG, T. RICKABAUGH, S. SPECK and R. SUN, 2001 Function of Rta is essential for lytic replication of murine gammaherpesvirus 68. *J Virol* **75**: 9262-9273.
- YOUNG, L. S., J. R. ARRAND and P. G. MURRAY, 2007 EBV gene expression and regulation in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, edited by A. ARVIN, G. CAMPADELLI-FIUME, E. MOCARSKI, P. S. MOORE, B. ROIZMAN *et al.*, Cambridge.
- YUAN, J., R. ADAMSKI and J. CHEN, 2010 Focus on histone variant H2AX: to be or not to be. *FEBS Lett* **584**: 3717-3724.
- ZHANG, F., L. CONG, S. LODATO, S. KOSURI, G. M. CHURCH *et al.*, 2011 Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29**: 149-153.
- ZU, Y., X. TONG, Z. WANG, D. LIU, R. PAN *et al.*, 2013 TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods* **10**: 329-331.
- ZUR HAUSEN, H., 2012 Red meat consumption and cancer: reasons to suspect involvement of bovine infectious factors in colorectal cancer. *Int J Cancer* **130**: 2475-2483.

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES  
DER FAKULTÄT FÜR BIOLOGIE  
DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

# TALE NUCLEASES AS ANTIVIRAL FACTORS

## Appendix

**MARKUS JULIUS PRÄGERT**  
MÜNCHEN 2013

## 10. Appendix

### 10.1 Oligonucleotides

#### 10.1.1 Mutagenesis PCR primers

**Table A1** PCR primer pairs used to introduce point mutations

FokI name	mutation	forward	backward	based on
RR	D483R	5'-Pho- CGGGAAATGCAACGATA TGTCGAAGAAAATC	TGCTTGGCCAATTGGCAG ATT	-
DD	R487D	5'-Pho- GACTATGTCTGAAGAAAA TCAAACACGAAAC	TTGCATTTTCATCTGCTTG GCC	-
-	E490K	5'-Pho- AAGGAAAATCAAACACG AAACAAACATATCAAC	GACATATCGTTGCATTTTC ATCTGCTTG	-
KK	I538K	5'-Pho- AAGACTAATTGTAATGG AGCTGTTCTTAGTGTAG AAGAG	ATGATTTAATCGTGTA GCTGAGCTTTGTAG	E490K
KKR	H537R	AAGACTAATTGTAATGG AGCTGTTC	5'-Pho- CCGATTTAATCGTGTA GCTGAGC	E490K I538K
-	Q486E	5'-Pho- GAGCGATATGTCTGAAGA AAATCAAACACG	CATTTTCATCTGCTTGGCC AATTG	-
EL	I499L	5'-Pho- CTGAACCCTAATGAATG GTGGAAAGTCTATCC	ATGTTTGTTTCGTGTTTG ATTTTCTTCG	Q486E
ELD	N496D	5'-Pho- GACAAACATCTGAACCC TAATGAATGG	TCGTGTTTGATTTTCTTC GACATATCG	Q486E I499L
inactivated cleavage domain	D450A	5'-Pho- GATTACGGTGTGATCGT GGCTACTAAAGC	5'-Pho- AATAGGAGATCCGACAG TATAAATTGCTC	-

-	HAX3 sequence in <i>gfp</i>	5'-Pho- TATCTCTAAGCTGGAGT ACAAC TACAACAGCC	AAGGGATTTCAGCTCGAT GCGGTTCAC	-
---	-----------------------------------	--	---------------------------------	---

### 10.1.2 PCR primers

**Table A2** PCR primer pairs

name	forward	backward
in vitro cut <i>gfp</i> template	GCAGTACATCAATGGGCGTGG	ATCCTGCTCCTCCACCTCCG
EMSA PCR template	ACATGAAGCAGCAGCACTTC	TTCACCAGGGTGTCGCCCTCGAAC

### 10.1.3 qPCR primers

**Table A3** Primer pairs for qPCR

name	forward	backward
eGFP TALEN7&8 cleavage site	CGACTTCAAGGAGGACGGCAAC	GGGGTGTTCTGCTGGTAGTGGTC
MHV-68 ORF6	AGGCTTGACCGTGGAACCTG	GGGGAATCTTGTTACTGTTGTGAAAG

### 10.1.4 EMSA DNA templates

**Table A4** Oligonucleotides used as DNA templates in EMSA experiments

name	upper strand	lower strand
ZFN 1 and 2	AGCGCACCATCTTCTTCAAGGACGAC GGCAACTA	TAGTTGCCGTCGTCCTTGAAGAAGATGG TGCGCT
TALE 5	TGGAGTACAAC TACAACAGCCACAAC GTCTATATCATGGCCGACAAGCAGAA GAACGGCA	TGCCGTTCTTCTGCTTGTCGGCCATGAT ATAGACGTTGTGGCTGTTGTAGTTGTAC TCCA
HAX 3	TCATACGTGTCCCTTTATCTCTCTCA GCGCG	CGCGCTGAGAGAGATAAAGGGACACGTA TGA



GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGAC  
 GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG  
 ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCACCC  
 TGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA  
 AGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACT  
 ACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGG  
 GCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCC  
ACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTCAAGATCCGCC  
 ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCG  
 ACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCC  
 CAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGG  
 CATGGACGAGC

**Supplementary Figure S1** PCR product of the *gfp* sequence used as a DNA template in EMSAs with ZFN 2 and TALE 5. Target sequences of ZFN 2 (9 bp) and TALE 5 (19 bp) are underlined.

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGAC  
 GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG  
 ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCACCC  
 TGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA  
 AGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACT  
 ACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAATC  
CCTTTATCTCTAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGC  
 AGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG  
 CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACC  
 ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCC  
 TGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGC

**Supplementary Figure S2** PCR product of the modified *gfp* sequence harboring the HAX3 target sequence (underlined) used as a DNA template in EMSAs with HAX3

## 10.2 Amino acid sequences

### 10.2.1 Zinc finger nucleases

MDYKDDDDKGSAAWSHPQFEKGGGSGGGSGGGSWSHQFEKGASATMGPKKKRKVGRSMG  
 YPYDVDPDYALESRPGERPFQCRICMRNFSTNQKLEVHTRTHTGEKPFQCRICMRNFSVRHNLQ  
 RHLRTHHTGEKPFQCRICMRNFSQHPNLTRHLKTHLRPGAAARALVKSELEEKSELRHKLKYV  
 PHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTK  
 AYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQ  
 LTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF\*

**Supplementary Figure S3** Amino acid sequence of ZFN 1

MDYKDDDDKGSAAWSHPQFEKGGGSGGGSGGGSWSHQFEKGASATMGPKKKRKVGRSMG  
 YPYDVDPDYALESRPGERPFQCRICMRNFSAPSKLDRHTRTHTGEKPFQCRICMRNFSLGENLR  
 RHLRTHHTGEKPFQCRICMRNFSGGNLGRHLKTHLRPGAAARALVKSELEEKSELRHKLKYV  
 PHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTK  
 AYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQ  
 LTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF\*

**Supplementary Figure S4** Amino acid sequence of ZFN 2

### 10.2.2 TAL effector / nucleases

N-terminus:

APRRRAAQPSDASPAAQVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQ  
 HPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLL  
 KIAKRGGVTAVEAVHAWRNALTGAPLN

C-terminus:

SIVAQLSRPDPALAALT

FokI domain:

GSQLVKSELEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGG  
SRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVY  
PSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFN  
NGEINF

Repeat domain TALE/N 5:

LTPQQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASHDGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPQQVVAIASHDGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPQQVVAIASNIGGKQALETQVQALLPVLCQAAG  
LTPEQVVAIASHDGGKQALETQVQALLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPQQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPQQVVAIASHDGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASHDGGKQALETQVQRLPVLCQAAG  
LTPQQVVAIASHDGGKQALETQVQRLPVLCQAAG  
LTPEQVVAIASNIGGKQALETQVQRLPVLCQAAG  
LTPQQVVAIASHDGGKQALETQVQALLPVLCQAAG  
LTPQQVVAIASNIGGRPALE

Repeat domain TALEN 6:

LTPQQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NNGGK**QALETVQALLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQALLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NNGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NNGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NNGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**HDGGK**QALETVQALLPVLCQAHG  
LTPQQVVAIAS**NI**GGRPAL

Repeat domain TALEN 7:

LTPQQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**HDGGK**QALETVQALLPVLCQAHG  
LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG  
LTPEQVVAIAS**NGGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NNGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NNGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NNGGK**QALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HDGGK**QALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NI**GGKQALETVQALLPVLCQAHG  
LTPQQVVAIAS**NI**GGRPAL

Repeat domain TALEN 8:

LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NN**GGKQALETVQALLPVLCQAHG  
LTPEQVVAIAS**NG**GGKQALETVQALLPVLCQAHG  
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG  
LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG  
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG  
LTPQQVVAIAS**NI**GGKQALETVQALLPVLCQAHG  
LTPQQVVAIAS**NG**GGKPALE

**Supplementary Figure S5** Amino acid sequences of TALE/Ns used in my thesis

## 11. Publications

Docheva, D., Padula, D., Popov, C., Weishaupt, P., **Prägener, M.**, Miosge, N., Hickel, R., Böcker, W., Clausen-Schaumann, H., Schieker, M. (2010). "Establishment of immortalized periodontal ligament progenitor cell line and its behavioural analysis on smooth and rough titanium surface." *Eur Cell Mater.* 14;19:228-41.

Alberton, P., Popov, C., **Prägener, M.**, Kohler, J., Shukunami, C., Schieker, M., Docheva, D. (2012). "Conversion of human bone marrow-derived mesenchymal stem cells into tendon progenitor cells by ectopic expression of scleraxis." *Stem Cells Dev.* 10;21(6):846-58.

# Acknowledgments

An dieser Stelle möchte ich mich bei allen bedanken, die mich während meiner Doktorarbeit unterstützt haben. Mein besonderer Dank geht an:

**Prof. Dr. Wolfgang Hammerschmidt**, der mir die Promotion erst ermöglichte. Vielen Dank für das entgegengebrachte Vertrauen, die hervorragende wissenschaftliche Betreuung und das große Interesse an meinem Projekt!

**Prof. Dr. Dirk Eick** für die offizielle Betreuung meiner Doktorarbeit seitens der Universität, die unterstützenden Worte im Thesis Committee und die Unterstützung der „Xcelligence“ Proliferationsexperimente.

**Prof. Dr. Heiko Adler** für die Kooperation bei den “MHV-68 Experimenten”, das Interesse an meinem Projekt und sein jederzeit offenes Ohr für Diskussionen.

**Prof. Dr. Thomas Lahaye**, der mit seiner Idee das Projekt erst ins Rollen brachte und damit meine Doktorarbeit maßgeblich beeinflusst hat. Vielen Dank auch für die unkomplizierte Zusammenarbeit bei der Herstellung der TALENs!

**Orlando de Lange** für die super Kooperation bei der Herstellung der TALENs. Die unkomplizierte und schnelle Zusammenarbeit hat maßgeblich zum Gelingen des Projekts beigetragen.

**Prof. Dr. Toni Cathomen** für die ZFNs und nicht zuletzt die vielen hilfreichen Tipps am Anfang des Projekts.

**Prof. Dr. Berit Jungnickel** danke ich für die Vermittlung der Promotionsstelle!

**Dagmar Pich** für die Einweisung am FACS und die andauernde Unterstützung bei den diversen Sorts und sämtlichen Fragen der Zellkultur.

Ein Dankeschön auch an **Rosie** für die Unterstützung in den bürokratischen Dingen des Laboralltags. Mein besonderer Dank geht an alle (ehemalige) Kollegen in der Hammerschmidt-Gruppe **Bergi, Romana, Montse, Anne, Marisa, Kathrin, Alex, Taka, Mika, Steffi, Lisa, Markus** und **Christine** für die tolle Atmosphäre und Unterstützung. Ich danke Euch!

Ich danke Meiner **Familie** für die bedingungslose Unterstützung und das aufgebrachte Verständnis in der schwierigen Zeit während der Doktorarbeit.

**Steffi**. Ohne Dich hätte ich es nicht geschafft!!!